



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 121415

TO: Minh-Tam Davis
Location: REM/3A24/3C18
Art Unit: 1642
Monday, May 17, 2004
Case Serial Number: 10/048046

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: (571)272-2527
paul.schulwitz@uspto.gov

Search Notes

Examiner Davis,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(571)272-2527

07/99

STIC-Biotech/ChemLib

121415

From: Davis, Minh-Tam
Sent: Thursday, May 06, 2004 12:57 PM
To: STIC-Biotech/ChemLib
Subject: Search request for 10/048046

Please search in commercial database, PGPUB, issued patent files, and interference:

- 1) SEQ ID NO:1, 2.
- 2) Oligomer search for SEQ ID NO:1,
- 3) Amino acids 31-103, 303-346, 476-641 of SEQ ID NO:2.

Thank you.

MINH TAM DAVIS
ART UNIT 1642
ROOM 3A24, MB 3C18
272-0830

RECEIVED
MAY - 6 2004
STIC-BIOTECH/ChemLib
(STIC)

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 5/17
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

Db	625136	AAGCGTACTATTATATACAAATTTATTTAAAAAGTAATAATTTACAATGGCCCTCTTCAT	625197
Qy	2524	ACACAATATCTGCTGCTGGGAAAAACCACAGCAATTTTATCTATTTTATTTTATTAATAGT	2583
Db	625198	GCAAAATATATGCAATGCGAAATGGAAGAACGACCTGCTCTTTCTTTTAAATATATTAATAAAA	625257
Qy	2584	TTGTGCTTATCTCTCTAATAAGATTTTAAATGTCACAAACTGTAGCACAAATTAATATAAT	2643
Db	625258	AAATTTATTTCTCTAAAAAAAAGAAAAACAGTCAGATAATTTTGGAAAAATAACAATAAAG	625317
Qy	2644	TATAATTTACAAATTTGACAAAAA	2667
Db	625318	ACAAAATAAAAAATAATAGATAA	625341

RESULT 15

```

US-09-293-322C-8
; Sequence 8, Application US/09293322C
; Patent No. 6232110
; GENERAL INFORMATION:
; APPLICANT: Pallas, David C
; APPLICANT: Du, Xianxing
; TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methylsterase.
; Patent No. 6232110
; TITLE OF INVENTION: Recombinant DNA Molecules and Methods
; FILE REFERENCE: 105-97
; CURRENT APPLICATION NUMBER: US/09/293,322C
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 60/082,202
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2409)
; OTHER INFORMATION: N is A, T, G or C.
US-09-293-322C-8

```

[illegible]

Search completed: May 14, 2004, 20:10:15
Job time : 228 secs

960

```

/ APPLICANT: OLEK, Alexander
/ APPLICANT: FIEPENBROCK, Christian
/ APPLICANT: BERLIN, Kurt
/ TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
/ TITLE OF INVENTION: by Assessing DNA Methylation
/ FILE REFERENCE: 5013.1012
/ CURRENT APPLICATION NUMBER: US/10/204,708
/ CURRENT FILING DATE: 2003-05-06
/ PRIOR APPLICATION NUMBER: PCT/EP01/03971
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: DE 10019058.8
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: DE 10019173.8
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: DE 10032529.7
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: DE 10043826.1
/ PRIOR FILING DATE: 2000-09-01
/ NUMBER OF SEQ ID NOS: 98
/ SEQ ID NO 62
/ LENGTH: 6801
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
/ US-10-204-708-62

```

Query Match	1.6%	Score 42.4	DB 4	Length 6801
Best Local Similarity	58.9%	Pred. No. 1.2		
Matches 73	Conservative 0	Mismatches 51	Indels 0	Gaps 0
QY	2556	ATTTTATCTATTTTATTTTAAATAGTTGGTGCTTATCTCTTAATAAGATTTAAATGT	2615	
DB	2997	ATAAATCCAACTTTTTCTTTTATTACTATTACTTTTAAATATATAAATTTAAATA	2938	
QY	2616	CACAAACGTGAGCACAAATTAATAATTTATAAATTACAAATTCACAAAAAATAAAAA	2675	
DB	2937	ATTTCACCTAACAACAATTAATAATAATTAATAAAAAAATACTAAAAAATAAATA	2878	
QY	2676	AAAA	2679	
DB	2877	AAAA	2874	

RESULT 14
 US-09-790-988-1
 ; Sequence 1, Application US/09790988
 ; Patent No. 6632935
 ; GENERAL INFORMATION:
 ; APPLICANT: SHIGENOBU, SHUJI
 ; APPLICANT: WATANABE, HIDEMI
 ; APPLICANT: MATTORI, MASAHIRA
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
 ; FILE REFERENCE: 081356/0159
 ; CURRENT APPLICATION NUMBER: US/09/790,988
 ; CURRENT FILING DATE: 2001-02-23
 ; PRIOR APPLICATION NUMBER: JP2000-107160
 ; PRIOR FILING DATE: 2000-04-07
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 640681
 ; TYPE: DNA
 ; ORGANISM: Buchnera sp.
 US-09-790-988-1

Query Match	1.6%	Score 42.4;	DB 4;	Length 640681;
Best Local Similarity	50.5%;	Pred. No.13;		
Matches 103;	Conservative	0;	Mismatches 101;	Indels 0;
Gaps	0;			

2464	AAGTTTATCAAAAACATTGTTTCAGGAGAGGAGCATAACTTTACAGCCCTACAGGACGT	2533

QY	
----	--

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: FIBRO101
; CLONE: 53219
; US-08-910-925-2

Query Match      1.6%; Score 42.6; DB 3; Length 2369;
Best Local Similarity 58.1%; Pred. No. 0.58;
Matches 75; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 1639 TTGCGCACCTGACTGAGCTGACCGCGGCTGCTACGCTGCTGCGCCCGCTTT 1698
Db 1873 TGCTGCCACCTGACTGAGCTGCTACTGGAACCTGCTGCTACTTCCACTGC 1814

Qy 1699 TGTGAGCTCAACCTGGGTGACAACTGCTGACGCGCTGCTGACAAACACGCTACGAG 1758
Db 1813 TGGAACTACTTCCACTGCTGCTACTGTTGAACTGCTACTAGAACTGCTACTGCTACTGC 1754

Qy 1759 TCAGACATC 1767
Db 1753 TTCGACTTC 1745

RESULT 10
US-09-252-991A-2236/c
; Sequence 2236, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2236
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-2236

Query Match      1.6%; Score 42.4; DB 4; Length 966;
Best Local Similarity 54.5%; Pred. No. 0.4;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 74 CGGGGATGTGAATCCCGATGAGCGGCCGCGGAGGCAAGCAAGTCCGCGCGCGCAGC 133
Db 578 CGCTGCTGTTTCATGCCATCGATCGCTGACGACGACGACGAGCGCTGCGGAGG 519

Qy 134 CCTGGGACGCGCTCCTGCTGCTGGCGCGGAGGCGGAGCGCCACGCTCCTCTGAGGA 193
Db 518 CCGCTGGCGGGTCCGCGAGATGCGCGCGGAGGACCTGCCCGCGCGATCCTTGAGCGCG 459

Qy 194 AGCGGAGTGGACCATCGCGCGGAGACGAGGTTGCG 229
Db 458 AGCATCGGCTGCTGATCGTTCATGCGCGGTTGGCG 423

RESULT 11
US-09-252-991A-2607
; Sequence 2607, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; US-09-252-991A-2607

Query Match      1.6%; Score 42.4; DB 4; Length 1734;
Best Local Similarity 54.5%; Pred. No. 0.55;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 74 CGGGGATGTGAATCCCGATGAGCGGCCGCGGAGGCAAGCAAGTCCGCGCGCGCAGC 133
Db 1212 CGCTGCTGTTTCATGCCATCGATCGCTGACGACGACGAGCGCTTGGCGGAGG 1271

Qy 134 CCTGGGACGCGCTCCTGCTGCTGGCGCGGAGGCGGAGCGCCACGCTCCTCTGAGGA 193
Db 1272 CCGCTGGCGGGTCCGCGAGATGCGCGCGGAGATGCGCGCGCGGATGCTGCGCGCGG 1331

Qy 194 AGCGGAGTGGACCATCGCGCGGAGACGAGGTTGCG 229
Db 1332 AGCATCGGCTGCTGATCGTTCATGCGCGGTTGGCG 1367

RESULT 12
US-09-252-991A-2445
; Sequence 2445, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2445
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-2445

Query Match      1.6%; Score 42.4; DB 4; Length 2016;
Best Local Similarity 54.5%; Pred. No. 0.6;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 74 CGGGGATGTGAATCCCGATGAGCGGCCGCGGAGGCAAGCAAGTCCGCGCGCGCAGC 133
Db 1454 CGCTGCTGTTTCATGCCATCGATCGCTGACGACGACGAGCGCTGCGCGCGATCCTTGAGCGCG 1513

Qy 134 CCTGGGACGCGCTCCTGCTGCTGGCGCGGAGGCGGAGCGCCACGCTCCTCTGAGGA 193
Db 1514 CCGCTGGCGGGTCCGCGAGATGCGCGCGGAGATGCGCGCGGAGTCTGCGCGCGGATCCTTGAGCGCG 1573

Qy 194 AGCGGAGTGGACCATCGCGCGGAGACGAGGTTGCG 229
Db 1574 AGCATCGGCTGCTGATCGTTCATGCGCGGTTGGCG 1609

RESULT 13
US-10-204-708-62/c
; Sequence 62, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; CURRENT APPLICATION NUMBER: US/09/252,991A

```


Query Match	Score	DB 3	Length
Best Local Similarity	1.7%	45.4	991
Matches 106; Conservative	51.2%	0.068	
Matches 106; Conservative	0	0	0
Mismatches 101; Indels	0	0	0
Gaps	0	0	0

QY	2473	AAAAACATTTGTTTCAGAGAGCGGAGCATAGTTTTCAGCCCTACAGACGTTACACAATAT	2532
DB	766	AGAACCAATTTTAGAGGTTAGGATTTCATAAATAGCTTAGTATGATTTTGTGGGAAACAATTA	825

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 12:14:23 ; Search time 215 Seconds
(without alignments)
6914.946 Million cell updates/sec

Title: US-10-048-046-1

Perfect score: 2679

Sequence: 1 aagaattcgagcagggcg.....acaaaaaaaaaaaaaaaaa 2679

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgm2_6/ptodata/2/ina/5A COMB.seq:*

2: /cgm2_6/ptodata/2/ina/5B COMB.seq:*

3: /cgm2_6/ptodata/2/ina/6A COMB.seq:*

4: /cgm2_6/ptodata/2/ina/6B COMB.seq:*

5: /cgm2_6/ptodata/2/ina/PTCUS COMB.seq:*

6: /cgm2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51.4	1.9	7218	1	US-08-232-463-14
2	46.2	1.7	430	4	US-09-621-976-16656
3	45.8	1.7	1759	4	US-09-807-258-13
4	45.4	1.7	991	3	US-08-924-747-25
5	45.4	1.7	991	3	US-09-247-373B-25
6	45.4	1.7	991	3	US-09-296-715-25
7	45.2	1.7	4403765	3	US-09-103-840A-2
8	45.2	1.7	4411529	3	US-09-103-840A-1
9	42.6	1.6	2369	3	US-08-910-925-2
10	42.4	1.6	966	4	US-09-252-991A-2236
11	42.4	1.6	1734	4	US-09-252-991A-2607
12	42.4	1.6	2016	4	US-09-252-991A-2445
13	42.4	1.6	6801	4	US-10-204-708-62
14	42.4	1.6	640681	4	US-09-790-988-1
15	42.2	1.6	2409	3	US-09-293-322C-8
16	42.2	1.6	2409	4	US-09-839-97A-8
17	42.2	1.6	4255	5	PCT-US96-02331-14
18	42.2	1.6	4835	5	PCT-US96-02331-9
19	42.2	1.6	4835	5	PCT-US96-02331-9
20	41.4	1.5	2271	4	US-09-205-258-243
21	41.4	1.5	1037	4	US-09-489-847-112
22	41.2	1.5	8093	4	US-10-204-708-31
23	41.2	1.5	194	4	US-09-621-976-801
24	41	1.5	1661	4	US-09-409-096-5
25	40.8	1.5	1192	4	US-09-439-554-23
26	40.8	1.5	26000	4	US-09-843-376-10
27	40.8	1.5	26000	4	US-09-843-376-10

28	40.6	1.5	1637	4	US-09-205-258-178
29	40.4	1.5	370	4	US-09-376-113-1
30	40.4	1.5	505	4	US-09-621-976-15639
31	40.4	1.5	611	4	US-09-376-113-4
32	40.4	1.5	631	4	US-09-376-113-6
33	40.4	1.5	1114	4	US-09-152-060-41
34	40.4	1.5	1248	4	US-09-489-847-101
35	40.4	1.5	3572	4	US-09-575-574-3
36	40.4	1.5	1249	4	US-09-461-325-128
37	40.2	1.5	1249	4	US-10-012-542-128
38	40.2	1.5	1249	4	US-09-461-325-93
39	40.2	1.5	1260	4	US-10-012-542-93
40	40	1.5	850	4	US-09-252-991A-5753
41	40	1.5	837	4	US-09-252-991A-5702
42	40	1.5	1035	4	US-09-252-991A-5715
43	40	1.5	1245	4	US-09-252-991A-5725
44	40	1.5	1342	4	US-09-489-847-89
45	39.8	1.5	593	4	US-09-904-615-59

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 1.9%; Score 51.4; DB 1; Length 7218;

XX	The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention									
XX	SQ									
	Sequence	518 BP;	142 A;	127 C;	108 G;	136 T;	0 U;	5 Other;		
	Query Match	18.2%;	Score	488.8;	DB	4;	Length	518;		
	Best Local Similarity	98.8%;	Pred. No.	1.9e-107;						
	Matches	490;	Conservative	0;	Mismatches	6;	Indels	0;	Gaps	0;
Qy	2171	TTCA	CAGCCCTC	TAGGAGGAGG	CGCAGGGTCTCC	CACAGGTCTCTCGGGT	GACTCTT	2230		
Db	517	TTNN	CAGCCCTN	NAGGAGGAGG	CGCAGGGTCTCC	CACAGGTCTCTCGGGT	GACTCTT	458		
Qy	2231	CTGT	GGAGCTTTT	TACCTCTG	AGTGAGACCTCC	CAGAGCCCGGGGG	CGCAGGCCG	2290		
Db	457	CTGT	GGAGCTTTT	TACCTCTG	AGTGAGACCTCC	CAGAGCCCGGGGG	CGCAGGCCG	398		
Qy	2291	CCCT	CTGTG	AGCGCTGGG	CAGGGCTGCTGG	GTGGCATCAGCAGCAG	AGCCCTT	2350		
Db	397	CCCT	CTGTG	AGCGCTGGG	CAGGGCTGCTGG	GTGGCATCAGCAGCAG	AGCCCTT	338		
Qy	2351	CTGT	TAACATG	CGCGCGTCC	CGCAGAGGGG	CAGTTTGTCTTTT	GTACATTTCCGAA	2410		
Db	337	CTGT	TAACATG	CGCGCGTCC	CGCAGAGGGG	CAGTTTGTCTTTT	GTACATTTCCGAA	278		
Qy	2411	ACTA	CAGTTAAAG	CAGAGTCTGTTT	TTCAGGAAAGTTT	CAAGGAGAGAGG	CAAGTTTA	2470		
Db	277	ACTA	CAGTTAAAG	CAGAGTCTGTTT	TTCAGGAAAGTTT	CAAGGAGAGAGG	CAAGTTTA	218		
Qy	2471	TCAA	AACATTTGTT	TTCAGGAGAGG	AGCATAGTTTAC	AGCTTACAGGAG	GTACACAAT	2530		
Db	217	TCAA	AACATTTGTT	TTCAGGAGAGG	AGCATAGTTTAC	AGCTTACAGGAG	GTACACAAT	158		
Qy	2531	ATCT	GTGCTGGG	AAACACACAG	CAATTTATCTATTTT	TATTTTATTTT	TATAGTTTGGTC	2590		
Db	157	ATCT	GTGCTGGG	AAACACACAG	CAATTTATCTATTTT	TATTTTATTTT	TATAGTTTGGTC	98		
Qy	2591	TTAT	CTTCTAATA	AGATTTTAAATGT	CAAACTGTAGCA	CAAAATATATAT	TATTAATT	2650		
Db	97	TTAT	CTTCTAATA	AGATTTTAAATGT	CAAACTGTAGCA	CAAAATATATAT	TATTAATT	38		
Qy	2651	TACA	AATTGACAAAA	2666						
Db	37	TACA	AATTGACTAAAA	22						

Search completed: May 14, 2004, 12:44:57
Job time : 1498 secs

expressed sequence tag; EST; probe; chemotactic; proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal; antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian; antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic; cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine; autoimmune disorder; multiple sclerosis; allergic condition; ulcer; insulin dependent diabetes; asthma; myeloid cell deficiency; stroke; lymphoid cell deficiency; burn; osteoporosis; osteoarthritis; central nervous system disorder; Alzheimer's disease; stroke; Parkinson's disease; Huntington's disease; coagulation disorder; haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour; infection; depression; psoriasis; ss.

Homo sapiens.

WO200021991-A1.

20-APR-2000.

15-OCT-1999; 99WO-US024206.

15-OCT-1998; 98US-0104436P.

(GEM) GENETICS INST INC.

Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C; Merberg D, Treacy M, Bowman MR;

WPI; 2000-317938/27.

Isolated polynucleotides, and encoded proteins, comprising secreted expressed sequence tags (sESTs), useful for treating various disorders such as autoimmune, infectious, and central nervous system disorders.

Claim 1; Page 437; 803pp; English.

AAA43426 to AAA45925 represent specifically claimed secreted expressed sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue sources. The sESTs can have a range of activities depending on the tissues they were isolated from. The activities include: chemotactic; proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial; antiulcer; antiviral; antidiabetic; antiasthmatic; vulnary; antipsoriatic; osteopathic; neuroprotective; nootropic; antiparkinsonian; antipsoriatic; cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be used for gene therapy and in vaccines. The sESTs are useful as probes for the identification and isolation of full-length cDNAs and genomic DNA molecules which correspond to the sESTs. Proteins encoded by the sESTs are useful in assays for determining biological activity and raising antibodies. They may be useful for treatment of autoimmune disorders (multiple sclerosis, insulin dependent diabetes), allergic conditions (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers, osteoporosis, osteoarthritis, central nervous system disorders (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's disease), tumours, bacterial fungal or viral infections, depression and psoriasis. AAA45926 to AAA45931 represent linker variants which are given in the exemplification of the present invention

Sequence 575 BP; 124 A; 168 C; 165 G; 118 T; 0 U; 0 Other;

Query Match 18.5%; Score 495.8; DB 3; Length 575;

Best Local Similarity 94.8%; Pred No. 4e-109;

Matches 532; Conservative 0; Mismatches 7; Indels 22; Gaps 1;

743 GTGTCCTCTGTGCAAGTGTGAGTCTCCAGCTTTCAGCTTCTCCAGACAGAA 802

130 GTGTCCTCTGTGCAAGTGTGAGTCTCCAGCTTTCAGCTTCTCCAGACAGAA 189

803 AGACTGGCTCTCTTTTCGTTGGAAACCCAGGATCAGGAGATTTGGAGCCCGTGAAGA 862

130 AGACTGGCTCTCTTTTCGTTGGAAACCCAGGATCAGGAGATTTGGAGCCCGTGAAGA 249

863 AGAAATGAGAGG-----AGATGGGACCTTGGACCTGAACGGG 900

250 AGAAATGAGAGGAGAAACATCCCTTTCCTCTGTAGATGGGACCTTGGACCTGAACGGG 309

901 CAGTTGTTGGTGGCAACACCCGCTAGAAATGCCAAACCCGCTCCAGGAGCGTCAGAGCA 960

310 CAGTTGTTGGTGGCAACACCCGCTAGAAATGCCAAACCCGCTCCAGGAGCGTCAGAGCA 369

961 GCGGCTGGAAAGCCAGACCAAGATGAGAGGAGCGTGCATGCATCATCTGCCAGGACCTG 1020

370 GCGGCTGGAAAGCCAGACCAAGATGAGAGGAGCGTGCATGCATCATCTGCCAGGACCTG 429

1021 CTGCACGACTCGGTGAGTTTGCAGCCCTGCATGCACACAGTTCTCGCGGGTTGTACTCG 1080

430 CTGCACGACTCGGTGAGTTTGCAGCCCTGCATGCACACAGTTCTCGCGGGTTGTACTCG 489

1081 GCGTGAATGAGAGCGCTCGTCCCTGTCTCTACCTGCGCGCTCCCGTGGAGCGGATCTGT 1140

490 GCGTGAATGAGAGCGCTCGTCCCTGTCTCTACCTGCGCGCTCCCGTGGAGCGGATCTGT 549

1141 AAAAACCCACATCTCCAAAC 1161

550 AAAAACCCACATCTCCAAAC 570

RESULT 15

AAH11859/c

ID AAH11859 standard; cDNA; 518 BP.

XX AC AAH11859;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (3'-primer) SEQ ID NO:8694.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-00116126.

XX PR 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183767.

XX PR 09-JUN-2000; 2000JP-00241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX PS Claim 3; SEQ ID NO 8694; 2537pp + Sequence Listing; English.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0215647P.
PR 07-JUL-2000; 2000US-0215880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0225868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249292P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2003-147444/14.
XX P-PSDB; ABUS5385.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX Claim 1; SEQ ID NO 483; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.

CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 693 BP; 145 A; 214 C; 197 G; 123 T; 0 U; 14 Other;
Query Match 20.0%; Score 536.6; DB 7; Length 693;
Best Local Similarity 92.5%; Pred. No. 6.5e-119;
Matches 608; Conservative 7; Mismatches 13; Indels 29; Gaps 4;
Qy 1044 GCCTGCATGCACACGTTCTGCGCGCTTGTCTACTCGGGCTGATGGAGCGCTGCTCCCT 1103
Db 14 GCCTGCATGCACACGTTCTGCGCGCTTGTCTACTCGGGCTGATGGAGCGCTGCTCCCT 73
Qy 1104 GTGTCTTACCTGCGCTGCTCCGCGAGCGGATCTGTAAACACCATCTCTCAACACCT 1163
Db 74 GTGTCTTACCTGCGCTGCTCCGCGAGCGGATCTGTAAACACCATCTCTCAACACCT 133
Qy 1164 CGTGGAGCATATCCTCATCCAGCATCCAGAACAGTTCGAGTGAAGAGATGTCAAAG 1223
Db 134 CGTGGAGCATATCCTCATCCAGCATCCAGAACAGTTCGAGTGAAGAGATGTCAAAG 193
Qy 1224 TATGATGCCAGGAATAAATCACTAAGCATCTGCAGCCCAAGTTCAGGCGCTTTT 1283
Db 194 TATGATGCCAGGAATAAATCACTAAGCATCTGCAGCCCAAGTTCAGGCGCTTTT 253
Qy 1284 TTCTGATGAAGAGGAGTTTCAGAGGACCTGCTGGAGCTGTTCAGAGTTGACAGTGAATC 1343
Db 254 TTCTGATGAAGAGGAGTTTCAGAGGACCTGCTGGAGCTGTTCAGAGTTGACAGTGAATC 313
Qy 1344 CTCAGATTTAGCCAGCCATACGTTGTCGCGGAGTTCCTGAGTACAGAGGAGGCG 1403
Db 314 CTCAGATTTAGCCAGCCATACGTTGTCGCGGAGTTCCTGAGTACAGAGGAGGCG 373
Qy 1404 GGGCGACCTCCCTCCTGCGCCAGCACCCGAGGCGGAGCAGGAGCCACAGGCCCTGGG 1463
Db 374 GGGCGACCTCCCTCCTGCGCCAGCACCCGAGGCGGAGCAGGAGCCACAGGCCCTGGG 433
Qy 1464 GGATGCACCTCCACGCTCGTTCAGCCCTGACGACGAGTTCAGGATTAAGTGTGCGCTCT 1523
Db 434 GGATGCACCTCCACGCTCGTTCAGCCCTGACGACGAGTTCAGGATTAAGTGTGCGCTCT 489
Qy 1524 GCAAGGAAGCCAGCCCTGTGCACTGCTCTCCAGCCCATGCCGACCGGAGAGCGGA 1583
Db 490 GCAAGGAAGCCAGCCCTGTGCACTGCTCTCCAGCCCATGCCGACCGGAGAGCGGA 548
Qy 1584 GCGGAGCAGGAGCCCGGCTGTCGCCCTCAGCAGTGTGCGTGTGCTGCGAGCTTCTG 1643
Db 549 -CGGAGCAGGAGCCCGGCTGTCGCCCTCAGCAGTGTGCGTGTGCTGCGAGCTTCTG 584
Qy 1644 CCACCTGTACTGGGCTGCACTCCGAGCCGCTGCTAGCGCTGCTGCGCCCGTTTG 1700
Db 585 CCACCTGTACTGGGCTGCACTCCGAGCCGCTGCTAGCTTGTGCGCCCGTTTG 641
RESULT 14
AAA44336
ID AAA44336 standard; cDNA; 575 BP.
XX
XX AAA44336;
XX
XX 21-AUG-2000 (first entry)
XX
XX Human secreted expressed sequence tag SEQ ID NO:911.
XX
XX Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;

08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249219P.
17-NOV-2000; 2000US-0249220P.
17-NOV-2000; 2000US-0249221P.
17-NOV-2000; 2000US-0249222P.
17-NOV-2000; 2000US-0249223P.
17-NOV-2000; 2000US-0249224P.
17-NOV-2000; 2000US-0249225P.
17-NOV-2000; 2000US-0249226P.
17-NOV-2000; 2000US-0249227P.
17-NOV-2000; 2000US-0249228P.
17-NOV-2000; 2000US-0249229P.
17-NOV-2000; 2000US-0249230P.
01-DEC-2000; 2000US-0250180P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-488783/53.
P-PSDB; AAU16317.
New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives.
Claim 1; SEQ ID NO 483; 980pp; English.
The invention relates to isolated nucleic acid molecules and their
encoded secreted proteins. The nucleic acids and proteins are used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
in diagnosing a pathological condition or susceptibility to a
pathological condition. Antibodies to the proteins can also be used in
alleviating symptoms associated with the disorders and in diagnostic
immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
(ELISA). Disorders which are diagnosed or treated include autoimmune
diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia,
nervous system disorders e.g. Alzheimer's disease, infections caused by
bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
and many other disorders listed in the specification. The polypeptides
can also be used to aid wound healing and epithelial cell proliferation,
to prevent skin aging due to sunburn, to maintain organs before
transplantation, for supporting cell culture of primary tissues, to
regenerate tissues and in chemotaxis. The polypeptides can also be used
as a food additive or preservative to increase or decrease storage
capabilities, fat content, lipid, protein, carbohydrate, vitamins,
minerals, cofactors and other nutritional components. The present
invention encodes a novel secreted protein of the invention. Note: The
sequence data for this patent did not form part of the printed

Query Match 20.0%; Score 536.6; DB 4; Length 693;

Best Local Similarity 92.5%; Pred. No. 6.5e-119;
Matches 508; Conservative 7; Mismatches 13; Indels 29; Gaps 4;
QY 1044 GCCCTGATGACACAGCTTCTGGCGGCTTCTACTCGGGCTGGATGAGCGCTCGTCCCT 1103
DB 14 GCCCTGATGACACAGCTTCTGGCGGCTTCTACTCGGGCTGGATGAGCGCTCGTCCCT 73
QY 1104 GTGTCTTACCTGCGGCTGCTCCGTTGGAGCGGATCTGTAAAAACACATCTCTCAACACCT 1163
DB 74 GTGTCTTACCTGCGGCTGCTCCGTTGGAGCGGATCTGTAAAAACACATCTCTCAACACCT 133
QY 1164 CGTGAAGCATACCTCATTCAGCATCCAGCATGTCAGCATGTCGAGTGAAGAAGATGTGCAAG 1223
DB 134 CGTGAAGCATACCTCATTCAGCATCCAGCATGTCAGCATGTCGAGTGAAGAAGATGTGCAAG 193
QY 1224 TATGATGCCAGGAATAAATCACTCAAGCATCTCTCAGCCCCAAAGTTCAGGCGGCTCTTT 1283
DB 194 TATGATGCCAGGAATAAATCACTCAAGCATCTCTCAGCCCCAAAGTTCAGGCGGCTCTTT 253
QY 1284 TTCTGATGAAGAAGGAGTTTCAGAGGACCTGCTGGAGCTGTTCAGCTTGAAGTGAAGTGC 1343
DB 254 TTCTGATGAAGAAGGAGTTTCAGAGGACCTGCTGGAGCTGTTCAGCTTGAAGTGAAGTGC 313
QY 1344 CTCAGACATTAGCCAGCATACCTGCTGCGGCGAGTCTCTCAGTACAGAGGCGGCTGGG 1403
DB 314 CTCAGACATTAGCCAGCATACCTGCTGCGGCGAGTCTCTCAGTACAGAGGCGGCTGGG 373
QY 1404 GCGCAGGCTCTCCCACTGCCCCAGCACCCGAGGCGAGCCAGAGCCCCACAGGCGCTGGG 1463
DB 374 GCGCAGGCTCTCCCACTGCCCCAGCACCCGAGGCGAGCCAGAGCCCCACAGGCGCTGGG 433
QY 1464 GATGACCTTCAGCTCGCTGCGTCCAGTACAGCATGTCAGGATTCAGTGTGCGCTCT 1523
DB 434 GATGACCTTCAGCTCGCTGCGTCCAGTACAGCATGTCAGGATTCAGTGTGCGCTCT 489
QY 1524 GCAAGGAAGCCACGCGCTGTGCACTGCTGCTTCCAGCCCCATGCCCCGAGAGCGGA 1583
DB 490 GCAAGGAAGCCACGCGCTGTGCACTGCTGCTTCCAGCCCCATGCCCCGAGAGCGGA 548
QY 1584 GCGCAGAGGAGACCCCGGCTGCGCCCTCAGCAGTGTGCGTCTGCTGCGAGCGCTTCTG 1643
DB 549 -CGCGAGCAGGACCCCGGCTGCGCCCTCAGCAGTGTGCGTCTGCTGCGAGCGCTTCTG 584
QY 1644 CCACCTGTACTGGGCTGCAACCCGAGCGGCTGCTGCTGCTGCGCTGCGCTGCGCTTGTG 1700
DB 585 CCACCTGTACTGGGCTGCAACCCGAGCGGCTGCTGCTGCTGCGCTGCGCTTGTG 641

RESULT 13

ABX73645
ID ABX73645 standard; DNA; 693 BP.

AC ABX73645;

DT 18-MAR-2003 (first entry)

XX Human novel polynucleotide #473.

XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;
XX muscular disorder; respiratory disease; reproductive disorder;
XX gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX hyperproliferative disorder; inflammatory disease; allergic reaction;
XX blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
XX haemostatic; antiarteriosclerotic.

OS Homo sapiens.

XX US2002132753-A1.

PN 19-SEP-2002.

XX 17-JAN-2001; 2001US-00764864.

Db 598 TCCCTCTGTGCAAGTGNATGAGTCTCCAGCTTTTCCTCAGCTCTCCAGACAGAAAGAC 657
Qy 807 TGCCTCTCTTTTCGTCGTGGAACCCCGAGATCAGAGAGATTTGAGCCCGTGAAGAAGAA 866
Db 658 TGCCTCTCTTTTCGTCGTGGAACCCCGAGATCAGAGAGATTTGAGCCCGTGAAGAAGAA 717
Qy 867 AATCAGAGGAGATGGGACCTTGACCTGAACGGGCGAGTTGTTGTCGACACACCGCGTAG 926
Db 718 AATCAGAGGAGATGGGACCTTGACCTGAACGGGCGAGTTGTTGTCGACACACCGCGTAG 777
Qy 927 AATGCCCCAAACCGTCCACGAGGAGCTCAGACGAGCGCTGGGAAGC 973
Db 778 AATGCCCCAAACCGTCCACGAGGAGCTCAGACGAGCGCTGGGAAGC 824

RESULT 12
AAS26304
ID AAS26304 standard; cDNA; 693 BP.
XX
AC AAS26304;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 483.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
FN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214888P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218220P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.

Db 911 GCCGCACTCAGGTGAAGAGCTCACCGCCATGAAATTCATCATATCTGTGACACACAA 970
Qy 2072 GGTTCAAAACTAAGCATCCAGAGCCCTCAGCAGAGCTTTGAGCACTGAGGTGAAGAG 2131
Db 971 GGTTCAAAACTAAGCATCCAGAGCCCTCAGCAGAGCTTTGAGCACTGAGGTGAAGAG 1030
Qy 2132 CGTGTCTTTTAAATACAGACAAAGCAGCAAGTGTCTTTTCAAGCCCTCAGGGAAG 2191
Db 1031 CGTGTCTTTTAAATACAGACAAAGCAGCAAGTGTCTTTTCAAGCCCTCAGGGAAG 1090
Qy 2192 GAGCGCAGGCTCCCGACAGGTCTCTGGGGTGAATCTTTCTGTGGAGCTTTTACCTCT 2251
Db 1091 GAGCGCAGGCTCCCGACAGGTCTCTGGGGTGAATCTTTCTGTGGAGCTTTTACCTCT 1148
Qy 2252 GAGTGACACCTCCCGACAGCCCGGGGGCGCAGCCCGCTCTGCTGAGCGCTGGGC 2311
Db 1149 GAGTGACACCTCCCGACAGCCCGGGGGCGCAGCCCGCTCTGCTGAGCGCTGGGC 1208
Qy 2312 AGGCTCGTGTGGCATCAGCAGCAGAGAGCAAGCTTTCTGTAACATGCGGCGCTCCCG 2371
Db 1209 AGGCTCGTGTGGCATCAGCAGCAGAGAGCAAGCTTTCTGTAACATGCGGCGCTCCCG 1268
Qy 2372 CCGAGAGGGCAGTTTGTCTTTTGTACATTTTCCGAACTACAGTTAAGCAGAGTC 2431
Db 1269 CCGAGAGGGCAGTTTGTCTTTTGTACATTTTCCGAACTACAGTTAAGCAGAGTC 1328
Qy 2432 TGTCTTTCAGGAAAGTTTCAAGGAGAGAGGCAAGTGTATCAAAAACATTTTTCAGGAG 2491
Db 1329 TGTCTTTCAGGAAAGTTTCAAGGAGAGAGGCAAGTGTATCAAAAACATTTTTCAGGAG 1388
Qy 2492 AAGGAGCATAAGTTTACAGCTACAGCAGTACACATATCTGCTGCTGGGAAACCA 2551
Db 1389 AAGGAGCATAAGTTTACAGCTACAGCAGTACACATATCTGCTGCTGGGAAACCA 1448
Qy 2552 CAGCATTTTATCTATTTTATTTTATTTTATAGTTTGTGCTTATCTTCTAATAAGATTAA 2611
Db 1449 CAGCATTTTATCTATTTTATTTTATTTTATAGTTTGTGCTTATCTTCTAATAAGATTAA 1508
Qy 2612 ATGTCACAACTGAGCACAATAATTAATTAATTTATTAATTTACAAATTTGACAAAA 2666
Db 1509 ATGTCACAACTGAGCACAATAATTAATTAATTTATTAATTTGACTAAAA 1563

RESULT 10
AAH06828
ID AAH06828 standard; cDNA; 816 BP.
XX
AC AAH06828;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:3663.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Negai K, Otsuki T;

XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
PT
XX
PS Claim 1; SEQ ID NO 3663; 2537bp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the CC complementary strand of a polynucleotide which comprises one of the 5602 CC nucleotide sequences defined in the specification, where the CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination CC of an oligonucleotide comprising a sequence complementary to the CC complementary strand of a polynucleotide which comprises a 5'-end CC sequence and an oligonucleotide comprising a sequence complementary to a CC polynucleotide which comprises a 3'-end sequence, where the CC oligonucleotide comprises at least 15 nucleotides and the combination CC of the 5'-end sequence/3'-end sequence is selected from those defined in the CC specification. The primer sets can be used in antisense therapy and in CC gene therapy. The primers are useful for synthesizing polynucleotides, CC particularly full-length cDNAs. The primers are also useful for the CC detection and/or diagnosis of the abnormality of the proteins encoded by CC the full-length cDNAs. The primers allow obtaining of the full-length CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent CC oligonucleotides, all of which are used in the exemplification of the CC present invention
XX
SQ Sequence 816 BP; 182 A; 216 C; 253 G; 162 T; 0 U; 3 Other;
Query Match 27.1%; Score 727; DB 4; Length 816;
Best Local Similarity 93.5%; Pred. No. 9.4e-165;
Matches 793; Conservative 0; Mismatches 18; Indels 37; Gaps 2;
Qy 17 GCGCAATGTCTCTTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 76
Db 5 GCGCAATGTCTCTTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 64
Qy 77 GGATGTGAATCCGATGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 136
Db 65 GGATGTGAATCCGATGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 124
Qy 137 GGGGACGGCTCTCGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 196
Db 125 GGGGACGGCTCTCGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184
Qy 197 GGGAGTGGACCATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 256
Db 185 GGGAGTGGACCATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 244
Qy 257 TCTCTGGAGATCACTGTAGAATTGTAGTGGATGAAATTCAGGTGAGTGACCTGGAG 316
Db 245 TCTCTGGAGATCACTGTAGAATTGTAGTGGATGAAATTCAGGTGAGTGACCTGGAG 304
Qy 317 ATACGACGACCATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 376
Db 305 ATACGACGACCATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 364
Qy 377 CTTTACAGACTGGGATGTCATCTACTTGGTGTACAGGAAGTGAACCGGAACCAAG 436
Db 365 CTTTACAGACTGGGATGTCATCTACTTGGTGTACAGGAAGTGAACCGGAACCAAG 424
Qy 437 TGGCATACCTCTATGAATCTTTAAGTGAAGCAAGGATGACACAGAAATCTTTGAAG 496
Db 425 TGGCATACCTCTATGAATCTTTAAGTGAAGCAAGGATGACACAGAAATCTTTT--- 480
Qy 497 CTAACAAAGGAAATGTGTTCATGGACCAAAGATACCTCAGGTGAGTGAGGCGGAG 556

Db	383	-----	382
Qy	596	AGGTGTCTTTTGGAGAACCAAGCCATCAACATCGAGCTCAGACCTTCTCCCAAGCCT	555
Db	383	-----	382
Qy	656	CGGCTTTCCACGGAGCCTTCTCTGAGGGGAGAGCGTTCTCCAGTTGTGGGTCTG	715
Db	383	-----	388
Qy	716	GGGTTGGTGGCATCTCCCTTAAAGGAAGTGGTCCCTCTGTGGCAAGTGATGAAGTCTCCA	775
Db	389	GGGTTGGTGGCATCTCCCTTAAAGGAAGTGGTCCCTCTGTGGCAAGTGATGAAGTCTCCA	448
Qy	776	GCCTTGGCTCAGCTCTCCCAAGCAAGAAAGACTGGTCTCTTTTGGTGTGGAACCCCAAG	835
Db	449	GCCTTGGCTCAGCTCTCCCAAGCAAGAAAGACTGGTCTCTTTTGGTGTGGAACCCCAAG	508
Qy	836	ATCAGGAGGATTTGGAGCCCGTGAAGAAATAGAGAGGAGATGGGGACCTTGACCTGA	895
Db	509	ATCAGGAGGATTTGGAGCCCGTGAAGAAATAGAGAGGAGATGGGGACCTTGACCTGA	568
Qy	896	ACGGGCAAGTTTGGTGGCAACACCGGTAGAAATGCCAAAACGCTCCACGAGAGAGCTCA	955
Db	569	ACGGGCAAGTTTGGTGGCAACACCGGTAGAAATGCCAAAACGCTCCACGAGAGAGCTCA	628
Qy	956	GAGCAGCGGCTGGGAACCCAGACAAGATGAGAGGAGAGCTGACATGCACTATCTGCAGG	1015
Db	629	GAGCAGCGGCTGGGAACCCAGACAAGATGAGAGGAGAGAGCTGACATGCACTATCTGCAGG	688
Qy	1016	ACCTGCTGCACTGCGTGGATTTGACAGCCCTGTCATGCAACGTTCTGGCGCGCTTGCT	1075
Db	689	ACCTGCTGCACTGCGTGGATTTGACAGCCCTGTCATGCAACGTTCTGGCGCGCTTGCT	748
Qy	1076	ACTCGGCTGGATGGAGCGCTGCTCCCTGTGTCTACTCTCGCGCTGTCGGTGGAGCGGA	1135
Db	749	ACTCGGCTGGATGGAGCGCTGCTCCCTGTGTCTACTCTCGCGCTGTCGGTGGAGCGGA	808
Qy	1136	TCTGTAAAAACCAATCCTCAACAACTCTGTGAAGCATACCTCATCCAGCATCCAGACA	1195
Db	809	TCGTGTAAAAACCAATCCTCAACAACTCTGTGAAGCATACCTCATCCAGCATCCAGACA	868
Qy	1196	AGAGTCCAGTGAAGAGATGTGCAAGATATGGATGCCAGGATATAAATCACTCAAGACA	1255
Db	869	AGAGTCCAGTGAAGAGAGATGTGCAAGATATGGATGCCAGGATATAAATCACTCAAGACA	928
Qy	1256	TGCTGCAGCCCAAAGTCAGGCGGTCTTTTCTGATGAAGAGGAGATTACAGAGACCTGC	1315
Db	929	TGCTGCAGCCCAAAGTCAGGCGGTCTTTTCTGATGAAGAGGAGATTACAGAGACCTGC	988
Qy	1316	TGGAGCTGTGAGAGCTTGAAGTCAAGTTCCTCAGACATTAGCCAGCATAGTGTGTGTGC	1375
Db	989	TGGAGCTGTGAGAGCTTGAAGTCAAGTTCCTCAGACATTAGCCAGCATAGTGTGTGTGC	1048
Qy	1376	GGCAGTGTCTGAGTACAGAGGAGCGGGCGCAGCCTCCCTCACTGCCACAGCACCCGAGG	1435
Db	1049	GGCAGTGTCTGAGTACAGAGGAGCGGGCGCAGCCTCCCTCACTGCCACAGCACCCGAGG	1108
Qy	1436	GCGAGCCAGGAGCCCAAGGCGCTTGGGGATGCAACCTCCAGTCGCTCAGCTGACGA	1495
Db	1109	GCGAGCCAGGAGCCCAAGGCGCTTGGGGATGCAACCTCCAGTCGCTCAGCTGACGA	1168
Qy	1496	CAGCAGTCCAGGATTAAGTGTGCCCTCTGCAAGGAAGCCAGCCCTGTGCACTGTCTGCT	1555
Db	1169	CAGT-----CAGGATTAAGTGTGCCCTCTGCAAGGAAGCCAGCCCTGTGCACTGTCTGCT	1224
Qy	1556	TCCAGCCCATGCCGACCGGAGAGCGGAGCCGAGCAGGAGCA--CCGCGTGTGCCCTTCAG	1614
Db	1225	T-CAGCCCATGCCGACCGGAGAGTGGNAACCGCAACAGGACCCCGCGTGTGCGSCCTTAA	1283
Qy	1615	CA 1616	
Db	1284	CA 1285	

RESULT 9	
ADA52592	
ID	ADA52592 standard; cDNA; 2186 BP.
XX	
AC	ADA52592;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Human coding sequence, SEQ ID 160.
XX	
KW	Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW	Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW	inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX	
OS	Homo sapiens.
XX	
FN	EPI293569-A2.
XX	
PD	19-MAR-2003.
XX	
PF	21-MAR-2002; 2002EP-00006586.
XX	
PR	14-SEP-2001; 2001JP-00328381.
PR	24-JAN-2002; 2002US-0350435P.
XX	
PA	(HELI-) HELIX RES INST.
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.
XX	
PI	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI	Yanamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX	
DR	WPI: 2003-395539/38.
DR	P-PSDB; ADA54231.
XX	
PT	New polynucleotides encoding full-length polypeptides, e.g. secretory
PT	and/or membrane proteins, useful for developing medicines for diseases in
PT	which the gene is involved, or as target molecules for gene therapy.
XX	
PS	Claim 1; SEQ ID NO 160; 205pp; English.
XX	
CC	The present invention relates to novel human secretory or membrane
CC	proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC	ADA54071). The coding sequences are useful in the gene therapy of
CC	diseases caused by abnormalities of the proteins, e.g. cancer,
CC	inflammatory diseases, osteoporosis or neurological disease.
XX	
SQ	Sequence 2186 BP; 493 A; 555 C; 539 G; 599 T; 0 U; 0 Other;
	Query Match 32.7%; Score 876.2; DB 7; Length 2186;
	Best Local Similarity 99.4%; Pred. No. 1.7e-200;
	Matches 890; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
QY	1772 AGATTACCTGGCAACAGAGGTTGCATGTGAAACAAACATGTTGACCGAGAGCCTCGTGG 1831
DB	671 AGATTACCTGGCAACAGAGGTTGCATGTGAAACAAACATGTTGACCGAGAGCCTCGTGG 730
QY	1832 CTCTCCAGCGGAGTGTTCTGCTGCTGATTACAGAGTCACGGAGAGACACCGTTCCTGT 1891
DB	731 CTCTCCAGCGGAGTGTTCTGCTGCTGATTACAGAGTCACGGAGAGACACCGTTCCTGT 790
QY	1892 GTTACTGCTGTGGCTCGCAGCTTCGGTAGCTGACCTATACAGTATCGGCAGAACATTC 1951
DB	791 GTTACTGCTGTGGCTCGCAGCTTCGGTAGCTGACCTATACAGTATCGGCAGAACATTC 850
QY	1952 CTGCTTCCGAGTTGCCAGTGGCCGTAACTCCCGTCTGACTGCTACTGGGCGCGTAAC 2011
DB	851 CTGCTTCCGAGTTGCCAGTGGCCGTAACTCCCGTCTGACTGCTACTGGGCGCGTAAC 910
QY	2012 GCOCGACTCAGGTGAAGCTCACACGCCATGAATTCATATCTGTGACACAGCAA 2071

XX AC ABX73184;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polynucleotide #12.
XX KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX OS Homo sapiens.
XX PN US2002132753-A1.
XX PD 19-SEP-2002.
XX PF 17-JAN-2001; 2001US-00764864.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217498P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 22-AUG-2000; 2000US-0225758P.
PR 30-AUG-2000; 2000US-0226868P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 08-SEP-2000; 2000US-0229513P.
PR 21-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 27-SEP-2000; 2000US-0234997P.
PR 29-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 01-NOV-2000; 2000US-0241809P.
PR 17-NOV-2000; 2000US-0244617P.
PR 08-DEC-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX Rosen CA, Ruben SM, Barash SC;
XX PI
XX WPI: 2003-147444/14.
DR P-PSDB; ABUS4924.
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 22; 402pp; English.
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 1311 BP; 306 A; 359 C; 399 G; 239 T; 0 U; 8 Other;
Query Match 35.3%; Score 945; DB 7; Length 1311;
Best Local Similarity 81.1%; Pred. No. 3.5e-217; Indels 282; Gaps 4;
Matches 1267; Conservative 6; Mismatches 7;
Qy 56 TCCGGGTTCCGCGCGGGGGGATGTGAATCCCGATGAGCGGCCGAGGAGGCAAGC 115
Db 5 TCCGGGTTCCGCGCGGGGGGATGTGAATCCCGATGAGCGGCCGAGGAGGCAAGC 64
Qy 116 AGTCGGCGCGCGCGAGCCCTGGGGACGGCTCTGGCTCTGGCGGGGAGGGGGGAGC 175
Db 65 AGTCGGCGCGCGCGAGCCCTGGGGACGGCTCTGGCTCTGGCGGGGAGGGGGGAGC 124
Qy 176 CGCACGTCTCTCTGAGGAAGCGGGAGTGGACCATCGGGGGGAGACGAGGTTGCGACCTTT 235
Db 125 CGCACGTCTCTCTGAGGAAGCGGGAGTGGACCATCGGGGGGAGACGAGGTTGCGACCTTT 184
Qy 236 CCTTCCCGAGCAATAACTGGTCTCTGGAGATCAGTGTAGAAATTTAGTGGATGAAAAT 295
Db 185 CCTTCCCGAGCAATAACTGGTCTCTGGAGATCAGTGTAGAAATTTAGTGGATGAAAAT 244
Qy 296 CAGTCAAGGTGACACTGGAAGATACCGACCAAGTGGAAACAGTGAATTAACAAGCTGAAG 355
Db 245 CAGTCAAGGTGACACTGGAAGATACCGACCAAGTGGAAACAGTGAATTAACAAGCTGAAG 304
Qy 356 TTGTTAAGAAAGAGACATGCCCTTTACAGACTGGGGATGTCTACTTCTGGTGACAGGA 415
Db 305 TTGTTAAGAAAGAGACATGCCCTTTACAGACTGGGGATGTCTACTTCTGGTGACAGGA 364
Qy 416 AGAATGAACCGGAACACACGCTGGCATACTCTATGAATCTTTAACTGAAAAGCAAGGCA 475
Db 365 AGAATGAACCGGAACACACGCTGGCATACTCTATGAATCTTTAACTGAAAAGCAAGGCA 382
Qy 476 TGACACAAGAAATCCTTTTGAAGCTTAACAAGAAAAATGTGTTCCATGGGACCAAGATACCT 535
Db 383 ----- 382
Qy 536 CAGGTGCAGGTGACGGGCGAGGGGCGGATCCCGGGTCCCTCCGTCGTCGCGCGCACCTC 595

diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 1: SEO ID NO 22: 980pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemoraxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence encodes a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed

Query Match 35.3%; Score 945; DB 4; Length 1311;
Best Local Similarity 81.1%; Pred. No. 3.5e-217;
Matches 1267; Conservative 6; Mismatches 7; Indels 282; Gaps 4

56	QC	TCGGGTTCGGCGCGGGCGGGGATGTGMAATCCCGATGGAGCGGCGGAGGAGCGAAGC	115
5	DB	TCGGGTTCGGCGCGGGCGGGGATGTGMAATCCCGATGGAGCGGCGGAGGAGCGAAGC	64
116	QC	AGTCGCGCGCGCGCGAGCGCTCTGGGAGCGGCTCTTCGCTCTGGCGCGGAGGAGGCGAGC	175
65	DB	AGTCGCGCGCGCGCGAGCGCTCTGGGAGCGGCTCTTCGCTCTGGCGCGGAGGAGGCGAGC	124
176	QC	CGCAGCTCTCTCTGAGGAAGCGGGAGTGGACCATCGGCGCGGAGCAGAGTTCGACCTTT	235
125	DB	CGCAGCTCTCTCTGAGGAAGCGGGAGTGGACCATCGGCGCGGAGCAGAGTTCGACCTTT	184
236	QC	CTTCTCCCAAGCAATAAACTGGTCTCTCGAGATCACTGTAGAAATGTAGTGGATGAAAAAT	295
185	DB	CTTCTCCCAAGCAATAAACTGGTCTCTCGAGATCACTGTAGAAATGTAGTGGATGAAAAAT	244
296	QC	CAGGTCAAGTGTACATCTGGNAGATACCAAGCACCAAGTGAACAGATGATTAAACAGCTGGAAG	355
245	DB	CAGGTCAAGTGTACATCTGGNAGATACCAAGCACCAAGTGAACAGATGATTAAACAGCTGGAAG	304
356	QC	TTGTTAAGAAGCAGACATGCCCTTTACAGACTGGGATGTCACTACTTGGTGTACAGA	415
305	DB	TTGTTAAGAARCAGACATGCCCTTTACAGACTGGGATGTCACTACTTGGTGTACAGA	364
416	QC	AGNATCAACCGAAACACACGCTGGCATACCTCTATGAATCTTTAAGTGAAGAACGAGGCA	475
365	DB	AGNATCAACCGAAACACACGCTGGCATACCTCTATGAATCTTTAAGTGAAGAACGAGGCA	382
476	QC	TGACACAGAATCTTTGAAGCTAACAGGAAAAATGTGTTCCATGGGACCAAGATACCT	535
383	DB	TGACACAGAATCTTTGAAGCTAACAGGAAAAATGTGTTCCATGGGACCAAGATACCT	382
536	QC	CAGGTCAAGTGCAGGCGAGGGGCGGATCCCGGGTCCCTCCGTCGTGCGCGCGCAGCT	595
383	DB	CAGGTCAAGTGCAGGCGAGGGGCGGATCCCGGGTCCCTCCGTCGTGCGCGCGCAGCT	382
596	QC	AGGTGTGCTTTGAGGAACCAAGCCATCAACATCGACGTGAGACCTCTTCCCAAGCCT	655
383	DB	AGGTGTGCTTTGAGGAACCAAGCCATCAACATCGACGTGAGACCTCTTCCCAAGCCT	382

RESULT 8
ABX73184
ID ABX

3184
ABX73184 standard; DNA; 1311 BP.

PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
DR P-PSDB; AAO08972.
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 8963; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2259 BP; 531 A; 612 C; 615 G; 501 T; 0 U; 0 Other;
XX
XX
XX Query Match 48.3%; Score 1294.2; DB 4; Length 2259;
XX Best Local Similarity 99.6%; Pred. No. 3.3e-301;
XX Matches 1308; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
XX
QY 1354 AGCCAGCCATACGTCTGTGCGGCGAGTGTCTGAGTACAGAGGCGAGCGCGAGCCT 1413
DB 1965 AGCCAGCCATACGTCTGTGCGGCGAGTGTCTGAGTACAGAGGCGAGCGCGAGCCT 1906
QY 1414 CCCACTGCCCCAGACCCGAGGCGAGCGAGGAGCCCGACAGCCCTGGGGGATGACCC 1473
DB 1905 CCCACTGCCCCAGACCCGAGGCGAGCGAGGAGCCCGACAGCCCTGGGGGATGACCC 1846
QY 1474 TCCAGCTCGGTACGCTGACGACGAGTCCAGGATTAGTGTCCTCTGCAAGGAGC 1533
DB 1845 TCCAGCTCGGTACGCTGACGACGAGTCCAGGATTAGTGTCCTCTGCAAGGAGC 1786
QY 1534 CACGCCCTGTGCACTGCTGCTTCCAGCCCATGCCCCGAGCGGAGCGGAGCGAGCAG 1593
DB 1785 CACGCCCTGTGCACTGCTGCTTCCAGCCCATGCCCCGAGCGGAGCGGAGCGAGCAG 1726
QY 1594 GACCGGCTGCGCCCTCAGCAGTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1653
DB 1725 GACCGGCTGCGCCCTCAGCAGTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1666
QY 1654 TGGGGCTGCACCCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1713
DB 1665 TGGGGCTGCACCCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1606
QY 1714 GTGACAGAGTGTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1773
DB 1605 GTGACAGAGTGTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1546
QY 1774 AATTACTGTCACACGAGGTTTGCATGGAAGAAACATGTTGACCGAGAGCCTCGTGCT 1833
DB 1545 AATTACTGTCACACGAGGTTTGCATGGAAGAAACATGTTGACCGAGAGCCTCGTGCT 1486
QY 1834 CTCACGCGGAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1893
DB 1485 CTCACGCGGAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1426
QY 1894 TACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1953
DB 1425 TACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1366
QY 1954 GCTTCCGAGTGTGCGAGTGGCGGTAACATCCCGTCTGCTGCTGCTGCTGCTGCTGCT 2013
DB 1365 GCTTCCGAGTGTGCGAGTGGCGGTAACATCCCGTCTGCTGCTGCTGCTGCTGCTGCT 1306

QY 2014 CGCACTCAGGTGAAGCTCACCAGCCATGAATCAATCATATCTGTGAACAGCAAGG 2073
DB 1305 CGCACTCAGGTGAAGCTCACCAGCCATGAATCAATCATATCTGTGAACAGCAAGG 1246
QY 2074 TTCAAAACTAAGCATCCAGAGCCCTGAGCAGCTTTCAGCACTGAGGTGAAGAGCG 2133
DB 1245 TTCAAAACTAAGCATCCAGAGCCCTGAGCAGCTTTCAGCACTGAGGTGAAGAGCG 1186
QY 2134 TGTTTTTAAATACAGACAGCAGCTCAAGGTGTTTTTACAGCCCTTGAAGGAGGG 2193
DB 1185 TGTTTTTAAATACAGACAGCAGCTCAAGGTGTTTTTACAGCCCTTGAAGGAGGG 1126
QY 2194 ACGCAGGCTCTCCGACAGAGTGTCTCTGGGTGACTTCTGTGAGCTTTTACCTCTGA 2253
DB 1125 ACGCAGGCTCTCCGACAGAGTGTCTCTGGGTGACTTCTGTGAGC--TTTACCTCTGA 1068
QY 2254 GTGAGACCTTCCCGACAGCCCGGGGCGCAGCCGCTCTCTGTGAGCGTGGGCGAG 2313
DB 1067 GTGAGACCTTCCCGACAGCCCGGGGCGCAGCCGCTCTCTGTGAGCGTGGGCGAG 1008
QY 2314 GGCTCGTGTGGCATCAGCAGCAGAGAGCGTCTTCTGTAACTGCGGCCGCTCCGCC 2373
DB 1007 GGCTCGTGTGGCATCAGCAGCAGAGAGCGTCTTCTGTAACTGCGGCCGCTCCGCC 948
QY 2374 GAGAGGCGCAGTTTTGCTCTTTGTACATTTTCCGAACTACAGTTTAAAGCAGAGTCTG 2433
DB 947 GAGAGGCGCAGTTTTGCTCTTTGTACATTTTCCGAACTACAGTTTAAAGCAGAGTCTG 888
QY 2434 TTTTCAGAAAAGTTTCAAGGAGAGGCAAGTTTATCAAAAACATTTGTTTCAAGAGAA 2493
DB 887 TTTTCAGAAAAGTTTCAAGGAGAGGCAAGTTTATCAAAAACATTTGTTTCAAGAGAA 828
QY 2494 GGGAGCATAGTTTACAGCCTACAGGAGCTACAAATATCTCTGCTGGGAAACACACA 2553
DB 827 GGGAGCATAGTTTACAGCCTACAGGAGCTACAAATATCTCTGCTGGGAAACACACA 768
QY 2554 GCATTTATCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2613
DB 767 GCATTTATCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 708
QY 2614 GTCAAACTGTAGCACAATATATATTTATTTATTTTATTTTATTTTATTTTATTTT 2666
DB 707 GTCAAACTGTAGCACAATATATTTATTTATTTTATTTTATTTTATTTTATTTTATTT 655

RESULT 7

AAS25843
ID AAS25843 standard; cDNA; 1311 BP.

XX AAS25843;

XX DT 07-NOV-2001 (first entry)

XX DE Human cDNA encoding a novel secreted protein, Seq ID 22.

XX KW Human; immunosuppressive; antiarthritic; SS; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.

XX OS Homo sapiens.

XX PN WO200155322-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001341.

1636	QY	CTTTCTGCACCTGTACTCTGGGCTGCACCCGGACCGGCTGTACGGCTGCCCTGGCCCG	1699
1518	Db	CTTTCTGCACCTGTACTCTGGGCTGCACCCGGACCGGCTGTACGGCTGCCCTGGCCCG	1577
1696	QY	TTTTTGTGAGCTCAACCTTGGGTGACAAGTGTCTGGA CGGCGTGTGAAACAACAACAGCTAC	1755
1578	Db	TTTTTGTGAGCTCAACCTTGGGTGACAAGTGTCTGGA CGGCGTGTGAAACAACAACAGCTAC	1637
1756	QY	GAGTCAGACATCCTGTGAAGATTACTTGGCAACACAGAGTTTGA CATGGA AAAACATGTTG	1815
1638	Db	GAGTCAGACATCCTGTGAAGATTACTTGGCAACACAGAGTTTGA CATGGA AAAACATGTTG	1697
1816	QY	ACCGAGAGCCTCTGTGGCTCTCCAGCGGGGAGTGTTCCTGCTGTCTGATTACAGAGTCACG	1875
1698	Db	ACCGAGAGCCTCTGTGGCTCTCCAGCGGGGAGTGTTCCTGCTGTCTGATTACAGAGTCACG	1757
1876	QY	GGAGACACCGTTCCTGTGTACTGTGTGTGGCTTCCGCGAGCTTCGCTGAGCTGACCTATCAG	1935
1758	Db	GGAGACACCGTTCCTGTGTACTGTGTGTGGCTTCCGCGAGCTTCGCTGAGCTGACCTATCAG	1815
1936	QY	TATCGGCAGAACATTCCTGCTTCCGAGTTCGCA GTGCGCTAAACATCCCGCTCTGACTGC	1995
1818	Db	TATCGGCAGAACATTCCTGCTTCCGAGTTCGCA GTGCGCTAAACATCCCGCTCTGACTGC	1877
1996	QY	TACTGGGGCGTAACCTGCGCACTCCAGGTGAAAGCTCACACAGCCCATGAAATTCATCAT	2055
1878	Db	TACTGGGGCGTAACCTGCGCACTCCAGGTGAAAGCTCACACAGCCCATGAAATTCATCAT	1937
2056	QY	ATCTGTGAAACAGACAAGAGTTCAAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTCAGCA	2115
1938	Db	ATCTGTGAAACAGACAAGAGTTCAAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTCAGCA	1997
2116	QY	CTGGAGTGAAGAGAGCGTGTTTTAAATATACAGACAAGCAGCTCAAGTGTTTTCAC	2175
1998	Db	CTGGAGTGAAGAGAGCGTGTTTTAAATATACAGACAAGCAGCTCAAGTGTTTTCAC	2057
2176	QY	AGCCCCCTGAGGGAAGGAGCGCAGGCTCCGACAGGTGTCTCTGGGTGACTCTCTGTGTG	2235
2058	Db	AGCCCCCTGAGGGAAGGAGCGCAGGCTCCGACAGGTGTCTCTGGGTGACTCTCTGTGTG	2117
2236	QY	GAGCTTTTACCCTCTGAGTGAGACCTCCCGAGAGCCCGGGGGCGCAGCCGCCCTC	2295
2118	Db	GAGC- -TTTACCCTCTGAGTGAGACCTCCCGAGAGCCCGGGGGCGCAGCCGCCCTC	2175
2296	QY	CTGCTGAGCGCTGGCGCAGGCTCGTGTGGCATCAGCAGCAGAGACGAAGCCTTTCTGTGA	2355
2176	Db	CTGCTGAGCGCTGGCGCAGGCTCGTGTGGCATCAGCAGCAGAGACGAAGCCTTTCTGTGA	2235
2356	QY	ACATGCCGCGTCCCGCCGAGAGGGCGAGTTTGTCTCTTTGTATCATTTTCCGAAACTAC	2415
2236	Db	ACATGCCGCGTCCCGCCGAGAGGGCGAGTTTGTCTCTTTGTATCATTTTCCGAAACTAC	2295
2416	QY	AGTTAAAGCAGAGTCTGTTTTTCAGGAAAAAGTTTCAAGGGAGAGGGCAAGTTTATCAAA	2475
2296	Db	AGTTAAAGCAGAGTCTGTTTTTCAGGAAAAAGTTTCAAGGGAGAGGGCAAGTTTATCAAA	2355
2476	QY	AACATTGTTTTCAGGAGAAGGGAGCATTAAGTTTACAGCCTACAGACGTCACATATCCT	2535
2356	Db	AACATTGTTTTCAGGAGAAGGGAGCATTAAGTTTACAGCCTACAGACGTCACATATCCT	2415
2536	QY	GCTCTCGGAAAAACACAGCATTTTATCTATTTTATTTTAAATAGTTTGGTGCTTATC	2595
2416	Db	GCCTCTGGAAAAACACAGCATTTTATCTATTTTATTTTAAATAGTTTGGTGCTTATC	2475
2596	QY	TTCTAAATGAATTTAAATGTGCACAACTGTGAGCAAAATATTAATTTTATTTTACAA	2655
2476	Db	TTCTAAATGAATTTAAATGTGCACAACTGTGAGCAAAATATTAATTTTATTTTACAA	2535
2656	QY	ATTGACAAAAA 2666	
2536	Db	ATTGACTAAAA 2546	

RESULT 5					
AAF89709	standard; DNA; 1995 BP.				
ID	AAF89709 standard; DNA; 1995 BP.				
XX					
AC	AAF89709;				
XX					
DT	22-AUG-2001 (first entry)				
XX					
DE	Nucleotide sequence of a human ring finger protein designated FHAR1..				
XX					
KW	FHAR1; RING finger protein; cancer; vaccine; ss.				
XX					
OS	Homo sapiens.				
XX					
FF	Key Location/Qualifiers				
FT	CDS 1..1995				
ET	/tag=a				
XX	/product= "ring finger protein FHAR1"				
FN	WO200142430-A1.				
XX					
PD	14-JUN-2001.				
XX					
PF	07-DEC-2000; 200WO-US033094.				
XX					
PR	08-DEC-1999; 99US-00456876.				
XX					
PA	(SWIK) SMITHKLINE BEECHAM CORP.				
XX					
PI	Zhou B, Zhu Y, Chaturvedi P, Hurle MR, Li X;				
XX					
DR	WPI; 2001-381663//40.				
DR	P-PSDB; AAB83843.				
XX					
PT	New FHAR1 polypeptide, a member of the RING finger protein family for				
PT	diagnosing and treating cancer, and for use in anti-cancer vaccines.				
XX					
PS	Claim 2; Page 18-19; 28pp; English.				
XX					
CC	The present sequence encodes a FHAR1 polypeptide, which is a member of				
CC	the RING finger protein family. FHAR1 is useful in the treatment of				
CC	cancer, and as a vaccine for inducing an immunological response in a				
CC	mammal. FHAR1 polynucleotides may also be used as a diagnostic reagent				
CC	through detection of mutations in the associated gene, and for chromosome				
CC	localization studies, and tissue expression studies. FHAR1 antibodies are				
CC	useful to isolate and to identify clones expressing the polypeptides, or				
CC	to purify the polypeptides by affinity chromatography and to treat cancer				
XX					
SQ	Sequence 1995 BP; 465 A; 563 C; 580 G; 387 T; 0 U; 0 Other;				
Query Match	74.5%; Score 1995; DB 5; Length 1995;				
Best Local Similarity	100.0%; Pred.No. 0;				
Matches 1995; Conservative	0; Mismatches 0; Indels 0; Gaps 0				
QY	91 ATGAGCGGCCGGAGGAGCAAGCAGTCGCCTCGCCGCGCCGAGCCCTGGGAGCGGTCTCTTG 150				
DB	1 ATGAGCGGCCGGAGGAGCAAGCAGTCGCCTCGCCGCGCCGAGCCCTGGGAGCGGTCTCTTG 60				
QY	151 CGTCTGGCGCGGAGAGGGCGAGCCGACGCTCTCTCTGAGMAGCGGAGTGGACCATC 210				
DB	61 CGTCTGGCGCGGAGAGGGCGAGCCGACGCTCTCTCTGAGMAGCGGAGTGGACCATC 120				
QY	211 GGGCGGAGACGAGGTGGGACTTTCTTCTTCCCAGCAATAAAGTGGTCTCTGGAGATCAC 270				
DB	121 GGGCGGAGACGAGGTGGGACTTTCTTCTTCCCAGCAATAAAGTGGTCTCTGGAGATCAC 180				
QY	271 TGTAGATTGTAGTGATGAATAATCAGGTGACGTGACCTGGAGATACCCAGCACCAGT 330				
DB	181 TGTAGATTGTAGTGATGAATAATCAGGTGACGTGACCTGGAGATACCCAGCACCAGT 240				
QY	331 GGAAACAGTGATTAAACAAGCTGAAGGTTGTTAAGAAGCAGACATGCCCTTTACAGACTGGG 390				
DB	241 GGAAACAGTGATTAAACAAGCTGAAGGTTGTTAAGAAGCAGACATGCCCTTTACAGACTGGG 300				

[illegible]

598	Db	TCCTCTGTGGCAAGTGAATGAAGTCTCCAGCTTTGCCCTCAGCTCTCTCCAGACAAGAAGAC	657
807	Qy	TGGGTCCCTTTTCGTCTCTTGAAACCCACAGGATCAGGAGGATTTGGAGCCCGTGAAGAAGAA	866
658	Db	TGGGTCCCTTTTCGTCTCTTGAAACCCACAGGATCAGGAGGATTTGGAGCCCGGAGAGAAGAA	717
867	Qy	AATGAGAGGAGATGGGACCTTGACTGAACGGGACGTTGTTGGTTCGCAACAACCGGTAG	926
718	Db	AATGAGAGGAGATGGGACCTTGACTGAACGGGACGTTGTTGGTTCGCAACAACCGGTAG	777
927	Qy	AAATGCCCAACCGTCCACGAGGACGTCAGAGACGGCTGGGAAGCCACGACAAGATGGA	986
778	Db	AAATGCCCAACCGTCCACGAGGACGTCAGAGACGGCTGGGAAGCCACGACAAGATGGA	837
987	Qy	GGAGACGCTGACATGCAATCATCTGCCAGGACCTGCTGCAAGCATGCTGTGAGTTTGACGCC	1046
838	Db	GGAGACGCTGACATGCAATCATCTGCCAGGACCTGCTGCAAGCATGCTGTGAGTTTGACGCC	897
1047	Qy	CTGCATGCAACAGTTTCTCGCGGCTTCTACTCGGGCTGGATGAGAGCGCTCGTCCCTGTG	1106
898	Db	CTGCATGCAACAGTTTCTCGCGGCTTCTACTCGGGCTGGATGAGAGCGCTCGTCCCTGTG	957
1107	Qy	TCCTACCTGCGCTGTCCGCTGGAGCGEATCTGTAAAAACCAATCTCTCAACACCTCGT	1166
958	Db	TCCTACCTGCGCTGTCCGCTGGAGCGEATCTGTAAAAACCAATCTCTCAACACCTCGT	1017
1167	Qy	GGAAAGCATACCTCATCAGCATCCAGACAAGATCGCAGTGAAGAAGATGTGCAAAGTAT	1226
1018	Db	GGAAAGCATACCTCATCAGCATCCAGACAAGATCGCAGTGAAGAAGATGTGCAAAGTAT	1077
1227	Qy	GGATGCCAGGAATAAAATCACTCAAGACATGCTGTCAGCCCAAAAGTCAGGCGGTCTTTTTC	1286
1078	Db	GGATGCCAGGAATAAAATCACTCAAGACATGCTGTCAGCCCAAAAGTCAGGCGGTCTTTTTC	1137
1287	Qy	TGATGAAGAAGGAGTTTCAGAGGACCTGCTGGAGCTGTCTAGAAGCTTGACAGTGAAGTCTC	1346
1138	Db	TGATGAAGAAGGAGTTTCAGAGGACCTGCTGGAGCTGTCTAGAAGCTTGACAGTGAAGTCTC	1197
1347	Qy	AGACATTTAGCCAGCATACGTCGTGTGTCGCGCAGTGTCTGAGTACAGAAGGACGAGGCGC	1406
1198	Db	AGACATTTAGCCAGCATACGTCGTGTGTCGCGCAGTGTCTGAGTACAGAAGGACGAGGCGC	1257
1407	Qy	GCAGCTTCCCACCTGCCCCAGCACCCGAGGGCGAGCCAGAGGCCCCACAGGCCCTGGGGGA	1466
1258	Db	GCAGCTTCCCACCTGCCCCAGCACCCGAGGGCGAGCCAGAGGCCCCACAGGCCCTGGGGGA	1317
1467	Qy	TGCACCTTCAAGTCCGTCTGAGCTGAAGACAGAGTCCAGGATTAAGTGTGCCCTCTGCA	1526
1318	Db	TGCACCTTCAAGTCCGTCTGAGCTGAAGACAGAGTCCAGGATTAAGTGTGCCCTCTGCA	1377
1527	Qy	AGGAAGCAGCCCTGTGCACCTGCTGCTTCCAGCCCATGCCCCAGCCGAGAGCGGAGCG	1586
1378	Db	AGGAAGCAGCCCTGTGCACCTGCTGCTTCCAGCCCATGCCCCAGCCGAGAGCGGAGCG	1437
1587	Qy	CGAGCAGGAACCGCGGTGTGCCCCCTCAAGAGTGTGCGGTCTGCCTGAGAGCCCTTCTGCCA	1646
1438	Db	CGAGCAGGAACCGCGGTGTGCCCCCTCAAGAGTGTGCGGTCTGCCTGAGAGCCCTTCTGCCA	1497
1647	Qy	CCTGTACTGGGGCTGCACCCGAGCCGCTGTCTAGCGGTCTGCCGCCCTTTTGTGAGCT	1706
1498	Db	CCTGTACTGGGGCTGCACCCGAGCCGCTGTCTAGCGGTCTGCCGCCCTTTTGTGAGCT	1557
1707	Qy	CAACTGGGTGAAGAATGTCTGGAACGGGTGCTGAAACAACAGCTACAGGTACAGAT	1766
1558	Db	CAACTGGGTGAAGAATGTCTGGAACGGGTGCTGAAACAACAGCTACAGGTACAGAT	1617
1767	Qy	CCTGAAGAAATTAACCTGCGAACCAAGGTTTGACATGGAAAAACAATGTTGACCGAGAGCCT	1826
1618	Db	CCTGAAGAAATTAACCTGCGAACCAAGGTTTGACATGGAAAAACAATGTTGACCGAGAGCCT	1677
1827	Qy	CGTGGCTCTCCAGCGGGAGTGTTTCTGCTGTCTGATTAAGAGTCAAGGAGACACCGT	1886

Db	1578	CGTGGCTCTCCAGGGGGAGTGTTTCTGCTGTCTGATTACAGAGTCCAGGGAGACACCGT	1733
Qy	1887	TCTGTGTTATCTGTGTGGCTCGGCAGCTTCCGTGAGCTGACCATCTACGTATCGGCAGAA	1946
Db	1738	TCTGTGTTATCTGTGTGGCTCGGCAGCTTCCGTGAGCTGACCATCTACGTATCAGCAGAA	1797
Qy	1847	CATTCTGCTTCCGAGTTGCCAGTGCGCGTAAACATCCGTCCTGACTGCTACTTGGGGCGG	2006
Db	1798	CATTCTGCTTCCGAGTTGCCAGTGCGCGTAAACATCCGTCCTGACTGCTACTTGGGGCGG	1857
Qy	2007	TAACTGCGGCACCTCAGGTGAAAGCTCACCACGCCATGMAATTCAATCATATCTGTGAACA	2066
Db	1858	TAACTGCGGCACCTCAGGTGAAAGCTCACCACGCCATGMAATTCAATCATATCTGTGAACA	1917
Qy	2067	GACAAGTTTCAAAAATAAGCATCCAGAGCCCTGAGAGCTTTACGACCTGGAGGTGAA	2128
Db	1918	GACAAGTTTCAAAAATAAGCATCCAGAGCCCTGAGAGCTTTACGACCTGGAGGTGAA	1977
Qy	2127	GAGAGCGTGTTTTTAAATACAGAGACAAGCAGTCAAGGTGTTTTTACAGAGCCCTGAG	2186
Db	1978	GAGAGCGTGTTTTTAAATACAGAGACAAGCAGTCAAGGTGTTTTTACAGAGCCCTGAG	2037
Qy	2187	GGAAAGGACACAGGGTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGGCTTTTTAC	2246
Db	2038	GGAAAGGACACAGGGTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGGCTTTTTAC	2095
Qy	2247	CTTCTGATGAGACCTTCCCGAGAGCCCCGGGGCGCGACGCCGCCCTCTGTGTAGCGCG	2306
Db	2096	CTTCTGATGAGACCTTCCCGAGAGCCCCGGGGCGCGACGCCGCCCTCTGTGTAGCGCG	2155
Qy	2307	TGGCAGGGCTCGTGGTGGCATCAGCAGCAGAGACGAGGCTTTCTGTAACTGCGCGCGG	2366
Db	2156	TGGCAGGGCTCGTGGTGGCATCAGCAGCAGAGACGAGGCTTTCTGTAACTGCGCGCGG	2215
Qy	2367	TCCCGCCGAGAGGGCGAGTTTGTCTCTTTTGTACATTTTCCGAAACTACAGTTAAAGCAG	2426
Db	2216	TCCCGCCGAGAGGGCGAGTTTGTCTCTTTTGTACATTTTCCGAAACTACAGTTAAAGCAG	2275
Qy	2427	AAGTCTGTTTTACGAAAAAGTTTCAAGGAGAAAGGGCAAGTTTATCAAAAACATTGTTTC	2486
Db	2276	AAGTCTGTTTTACGAAAAAGTTTCAAGGAGAAAGGGCAAGTTTATCAAAAACATTGTTTC	2335
Qy	2487	AGGAGAGGGAGCATAGTTTACAGCCTACAGAGCGTACACAATATCCTGCTGCTGGAA	2546
Db	2336	AGGAGAGGGAGCATAGTTTACAGCCTACAGAGCGTACACAATATCCTGCTGCTGGAA	2395
Qy	2547	AACACACAGCATTTATCTATTTTTTATTTTAAATAGGTTTGGTGCTTATCTTTAAATAAGA	2606
Db	2396	AACACACAGCATTTATCTATTTTTTATTTTAAATAGGTTTGGTGCTTATCTTTAAATAAGA	2455
Qy	2607	TTTAAATGTCAAAACTGTGAGCAAAATAATAATTTATAATTTACAAATTCAGCAAAA	2666
Db	2456	TTTAAATGTCAAAACTGTGAGCAAAATAATAATTTATAATTTACAAATTCAGCAAAA	2515

RESULT 4
ABN59646

ABN59646
ID ABN59646 standard: cDNA; 3181 BP.

AX
AC ABN59646;

28-JUN-2002 (first entry)

XX
DE
Novel human coding sequence
SEO ID NO: 57.

Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
 XX KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag; gene; ss.

XX Homo sapiens.

XX
PN W030022660-A2

XX PN

1217	QY	TGCAAAAGTAGTATGCAGGAATAAAATCACTCAAGACATGCTGCAGCCCAAAAGTCAGGC	1276
1169	DB	TGCAAAAGTAGTATGCAGGAATAAAATCACTCAAGACATGCTGCAGCCCAAAAGTCAGGC	1228
1277	QY	GGTCTTTTTCTGATGAAGAAGGAGTTCAAGAGCACTGCTGGAGCTGTCAAGCGTTTGACA	1336
1229	DB	GGTCTTTTTCTGATGAAGAAGGAGTTCAAGAGCACTGCTGGAGCTGTCAAGCGTTTGACA	1288
1337	QY	GTGAGTCTCAGACATTAAGCCAGCCCAATACGTGTGTGTCGGGCAAGTGTCTCTGAGTACAGAA	1396
1289	DB	GTGAGTCTCAGACATTAAGCCAGCCCAATACGTGTGTGTCGGGCAAGTGTCTCTGAGTACAGAA	1348
1397	QY	GGCAGCGGGCAGCCTCCCACTGCCCAAGCACCCCGAGGCGCAGCCAGAGAGCCCCACAGG	1456
1349	DB	GGCAGCGGGCAGCCTCCCACTGCCCAAGCACCCCGAGGCGCAGCCAGAGAGCCCCACAGG	1408
1457	QY	CCCTGGGGGATGCACCTCCACGTCGCTCAGCCTTGACGACAGCAGTCCAGGATTAACGTGT	1516
1409	DB	CCCTGGGGGATGCACCTCCACGTCGCTCAGCCTTGACGACAGCAGTCCAGGATTAACGTGT	1468
1517	QY	GCCTCTCAAGGAAGCAGCCCTGTGCACCTGCTGCTTCAGAGCCCATGCCGACCCGGA	1576
1469	DB	GCCTCTCAAGGAAGCAGCCCTGTGCACCTGCTGCTTCAGAGCCCATGCCGACCCGGA	1528
1577	QY	GAGCGAGGCGCAGCAGGACCCGCGTGTGCCCCCTCAGCAGTGTGCGGTCTGCCTGCAGC	1636
1529	DB	GAGCGAGGCGCAGCAGGACCCGCGTGTGCCCCCTCAGCAGTGTGCGGTCTGCCTGCAGC	1588
1637	QY	CTTTCTGCACCTGTACTGCGGCTGCACCGGACCGGCTGTACTACGCTGCCTGGCCCCGT	1696
1589	DB	CTTTCTGCACCTGTACTGCGGCTGCACCGGACCGGCTGTACTACGCTGCCTGGCCCCGT	1648
1697	QY	TTTGTGAGCTCAACCTGGGTGACAAGTGTCTGGACGGCGTCTGGAACAAACACAGACTACG	1756
1649	DB	TTTGTGAGCTCAACCTGGGTGACAAGTGTCTGGACGGCGTCTGGAACAAACACAGACTACG	1708
1757	QY	AGTCAGACATCTCTGAAGAAATTACCTGGCAACACAGAGTTTGACATGGAAAAACATGTTGA	1816
1709	DB	AGTCAGACATCTCTGAAGAAATTACCTGGCAACACAGAGTTTGACATGGAAAAACATGTTGA	1768
1817	QY	CCGAGAGCTCGTGGCTCTCCAGCGGGAGTGTTCGTCTGCTCTGATTACAGAGTACACGG	1876
1769	DB	CCGAGAGCTCGTGGCTCTCCAGCGGGAGTGTTCGTCTGCTCTGATTACAGAGTACACGG	1828
1877	QY	GAGACACCGTTCGTGTGTACTGTGTGGCTTGGCGAGCTTCGCTGAGCTGACCTATCAGT	1936
1829	DB	GAGACACCGTTCGTGTGTACTGTGTGGCTTGGCGAGCTTCGCTGAGCTGACCTATCAGT	1888
1937	QY	ATCGGAGAAACATTCCTGCTTCGAGTTGCGAGTGGCGGTAAACATCCCGTCTCTGACTGCT	1996
1889	DB	ATCGGAGAAACATTCCTGCTTCGAGTTGCGAGTGGCGGTAAACATCCCGTCTCTGACTGCT	1948
1997	QY	ACTGGGCGCTAACTCCGCACTCAGGTGAAAGCTCCACAGCCCATGAAATTCAAATCATTA	2056
1949	DB	ACTGGGCGCTAACTCCGCACTCAGGTGAAAGCTCCACAGCCCATGAAATTCAAATCATTA	2008
2057	QY	TTGTGAAACAGACAAGGTTCAAAAATAAGCATCCAGAGCCCTGAGAGCTTTTCAGCAC	2116
2009	DB	TTGTGAAACAGACAAGGTTCAAAAATAAGCATCCAGAGCCCTGAGAGCTTTTCAGCAC	2068
2117	QY	TGGAGTGAAGAGAGCGTGTTTTTTAAATAACAGAGCAAGCAGCTCAAGGTGTTTTTCACA	2176
2069	DB	TGGAGTGAAGAGAGCGTGTTTTTTAAATAACAGAGCAAGCAGCTCAAGGTGTTTTTCACA	2128
2177	QY	GCCTCTGAGGGGAAGGACCGAGGCTCCCGACAGTGTCTCTGGGTGACTCTTCTGTGG	2236
2129	DB	GCCTCTGAGGGGAAGGACCGAGGCTCCCGACAGTGTCTCTGGGTGACTCTTCTGTGG	2188
2237	QY	AGCTTTTAACTCTCAGTGTAGACCTCCCGAGAGCCCGGGGGCGCAGCCCGCCCTCC	2296
2189	DB	AGCTTTTAACTCTCAGTGTAGACCTCCCGAGAGCCCGGGGGCGCAGCCCGCCCTCC	2248
2297	QY	TGCTGAGCGCTGGCGAGGCTCGTGTGTGGCATCAGCAGCAGAGACGAAGCCTTCTGTGAA	2356

CC checkpoint protein Chfr (see AAB20219) having a forkhead associated
CC domain (FHA) and a ring finger domain. The protein is required for
CC regulation of the transition of cells from prophase to metaphase during
CC mitosis. It has ubiquitin-protein ligase activity. The Chfr checkpoint
CC was evident in primary human cells, but was inactivated in 4 of 8 human
CC cancer cell lines. In U2OS cells, a mutation was identified, which caused
CC a Val to Met amino acid substitution in the highly conserved C-terminal
CC Cys-rich region of the Chfr protein. In the absence of the Chfr
CC checkpoint cells subjected to mitotic stress condensed their chromosomes
CC despite failing to separate their chromosomes. Chfr may monitor
CC centrosome separation. Inactivation of the Chfr gene in human cancer is
CC theorized to underlie the increased sensitivity of cancer cells to
CC antimitotic drugs. Nucleic acids comprising the present sequence, or
CC sequences encoding at least amino acids 31-103, 303-346, 476-641 (or
CC their antisense sequences) are claimed. The Chfr cDNA was isolated from
CC an expressed sequence tag database sequence for cDNAs with FHA motifs.
CC Claimed methods of determining the tumorigenic potential of a cell
CC comprise examining the cell for the presence of Chfr nucleic acid
CC (absence or mutation indicating predisposition to tumorigenesis upon
CC exposure to mitotic stress). Also claimed is a diagnostic agent, e.g. an
CC antisense fragment of the present sequence, that binds to the Chfr
CC nucleic acid, and a diagnostic kit for detecting tumorigenic potential
CC of a cell. A composition which inhibits the biological activity of Chfr
CC may comprise a ligand selected from an antibody or its fragment. The Chfr
CC inhibitor is used in a claimed method of retarding the growth of a cancer
CC cells. Chfr polypeptides are useful in screening for drugs which can
CC inhibit the activity of Chfr in a cancer cell, rendering the cell more
CC sensitive to additional antitumour therapies

XX Sequence 2679 BP; 650 A; 715 C; 764 G; 550 T; 0 U; 0 Other;

Query Match 100.0%; Score 2679; DB 5; Length 2679;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGATTTCGGCAGAGCGCCCAATGCTCTTGACAGCGCGCGGCGAGCCGGTTCGCG 60
DB 1 AAGATTTCGGCAGAGCGCCCAATGCTCTTGACAGCGCGCGGCGAGCCGGTTCGCG 60
QY 61 GTTCGGCGCGCGCGGGGATGTGAATCCGATGAGCGCGCGCGGCGAGGCAAGCATGCG 120
DB 61 GTTCGGCGCGCGCGGGGATGTGAATCCGATGAGCGCGCGCGGCGAGGCAAGCATGCG 120
QY 121 CCGCGCGCGAGCCCTGGGACAGCTCTCTGCTGGGCGCGAGAGGCGGCGAGCGCGAC 180
DB 121 CCGCGCGCGAGCCCTGGGACAGCTCTCTGCTGGGCGCGAGAGGCGGCGAGCGCGAC 180
QY 181 GTCTCTCTGAGGAGCGGAGTGAGACCATCGGCGGAGAGAGGTTGCGACCTTCTCTTC 240
DB 181 GTCTCTCTGAGGAGCGGAGTGAGACCATCGGCGGAGAGAGGTTGCGACCTTCTCTTC 240
QY 241 CCGAGCAATAACTGGTCTCTGGAGATCACTGTAGATTTAGTGGATGAAATCAGGT 300
DB 241 CCGAGCAATAACTGGTCTCTGGAGATCACTGTAGATTTAGTGGATGAAATCAGGT 300
QY 301 CAGGTGACATGGAAGATACAGCACCAGTGGAACTGATTAACTGAAGCTGAAGTTGTT 360
DB 301 CAGGTGACATGGAAGATACAGCACCAGTGGAACTGATTAACTGAAGCTGAAGTTGTT 360
QY 361 AAGAGCAGACATCCCTTTACAGACTGGGGATGTCATCTACTTGGTATCAGGAGAT 420
DB 361 AAGAGCAGACATCCCTTTACAGACTGGGGATGTCATCTACTTGGTATCAGGAGAT 420
QY 421 GAACCGGAAACACACGTGGCATACCTCTATGAATCTTTTAAGTGAAGACGAGGCATGACA 480
DB 421 GAACCGGAAACACACGTGGCATACCTCTATGAATCTTTTAAGTGAAGACGAGGCATGACA 480
QY 481 CAAGAATCTTTGAAGCTAACAGAGAAATGTGTTCATGGGACCAAGATACCTCAGGT 540
DB 481 CAAGAATCTTTGAAGCTAACAGAGAAATGTGTTCATGGGACCAAGATACCTCAGGT 540
QY 541 GCAGGTGAGGCGGAGGCGGATCCCGGGTCCCTCCGCTCGCGCGCCCATCTCAGGTG 600

DB 541 GCAGGTGAGGCGGAGGCGGATCCCGGGTCCCTCCGTCGTCGCCGCCACTCAGGTG 600
QY 601 TGCTTTTCAGGAACACACAGCCCATCAATCAAGCTCAGACCTCTTTCCCCACAGCCTCGGCC 660
DB 601 TGCTTTTCAGGAACACACAGCCCATCAATCAAGCTCAGACCTCTTTCCCCACAGCCTCGGCC 660
QY 661 TCTTCCACGAGACCTTCTCTGACGAGGCGAGAGGCTTCTCCAGTTGTTGGTCTGGGGT 720
DB 661 TCTTCCACGAGACCTTCTCTGACGAGGCGAGAGGCTTCTTCCAGTTGTTGGTCTGGGGT 720
QY 721 GGTGCGATCTCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTGATGAAGTCTCCAGCTTT 780
DB 721 GGTGCGATCTCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTGATGAAGTCTCCAGCTTT 780
QY 781 GCCTCAGCTCTCCACAGAGAACTGCGTCTCTTTTCCTGCTGTTGGAAACCCAGGATCAG 840
DB 781 GCCTCAGCTCTCCACAGAGAACTGCGTCTCTTTTCCTGCTGTTGGAAACCCAGGATCAG 840
QY 901 CAGTTGTTGGTCGACAAACCCGCTAGAAATGCCAAACCCGCTCCACGAGGACGTGAGACGA 960
DB 901 CAGTTGTTGGTCGACAAACCCGCTAGAAATGCCAAACCCGCTCCACGAGGACGTGAGACGA 960
QY 961 CGGCTGGGAAGCCAGACAGATGCGGAGAGCGCTGACATGCATCTATGCCAGGACCTG 1020
DB 961 CGGCTGGGAAGCCAGACAGATGCGGAGAGCGCTGACATGCATCTATGCCAGGACCTG 1020
QY 1021 CTGCACGACTCGGTGAGTTTGCAGCCCTGCATGACACACGTTCTGCGGGGCTTGTACTCG 1080
DB 1021 CTGCACGACTCGGTGAGTTTGCAGCCCTGCATGACACACGTTCTGCGGGGCTTGTACTCG 1080
QY 1081 GGTGAGTGGAGCGCTCGTCCCTGTGCTTACCTTCCGCTGCTGCTGCTGAGCGGATCTGT 1140
DB 1081 GGTGAGTGGAGCGCTCGTCCCTGTGCTTACCTTCCGCTGCTGCTGCTGAGCGGATCTGT 1140
QY 1141 AAAAACACATCTCTCAACAACTCGTGAAGCATACCTCATCCAGCATCCAGACAAAGT 1200
DB 1141 AAAAACACATCTCTCAACAACTCGTGAAGCATACCTCATCCAGCATCCAGACAAAGT 1200
QY 1201 GCGAGTGAAGAGATGTGCAAGTATGGATGCCAGGAATAAAATCACTCAAGACATGCTG 1260
DB 1201 GCGAGTGAAGAGATGTGCAAGTATGGATGCCAGGAATAAAATCACTCAAGACATGCTG 1260
QY 1261 CAGCCCAAGTCAGCGCTCTTTTCTGATGAAGAAGGAGTTCAGAGGACCTGCTGGAG 1320
DB 1261 CAGCCCAAGTCAGCGCTCTTTTCTGATGAAGAAGGAGTTCAGAGGACCTGCTGGAG 1320
QY 1321 CTGTGAGACGTTGACAGTGAAGTCTCAGACATTAAGCCAGCCATACGTGCTGTGCGGAG 1380
DB 1321 CTGTGAGACGTTGACAGTGAAGTCTCAGACATTAAGCCAGCCATACGTGCTGTGCGGAG 1380
QY 1381 TGTCTGTGATCAGAAAGCAGCGCGGCGAGCTCCCACTGCCAGCACCCGAGGCGAG 1440
DB 1381 TGTCTGTGATCAGAAAGCAGCGCGGCGAGCTCCCACTGCCAGCACCCGAGGCGAG 1440
QY 1441 CCAGGAGCCCAACAGCCCTGGGGATGCAACCTCCACGCTCCGTCAGCTGACGACGA 1500
DB 1441 CCAGGAGCCCAACAGCCCTGGGGATGCAACCTCCACGCTCCGTCAGCTGACGACGA 1500
QY 1501 GTTCCAGATTAAGTGTGCTTCTGCAAGAGCCAGCCCTGTGCACTGCTGCTTCCAG 1560
DB 1501 GTTCCAGATTAAGTGTGCTTCTGCAAGAGCCAGCCCTGTGCACTGCTGCTTCCAG 1560
QY 1561 CCATGCGCGAGCCGAGAGCGGAGCGGAGAGGACCCGCGTGTGCGCCCTCAGCAGTGT 1620
DB 1561 CCATGCGCGAGCCGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGTGT 1620
QY 1621 GCGGTGCTGCTGACGCTTTCTGCACTGTGATCTGGGCTGTGACCCGAGCGGCTGTAC 1680
DB 1621 GCGGTGCTGCTGACGCTTTCTGCACTGTGATCTGGGCTGTGACCCGAGCGGCTGTAC 1680

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 11:05:18 ; Search time 1484 seconds
(without alignments)
7669.085 Million cell updates/sec

Title: US-10-048-046-1
Perfect score: 2679
Sequence: 1 aagaattcgcagagggcg.....acaaaaaaaaaaaaaaaaa 2679

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2679	100.0	2679	5	AAF30352 Human chf
2	2566.4	95.8	2639	4	AAH14542 Human CDN
3	2262.4	84.4	3138	4	AAH14556 Human CDN
4	2244.6	83.8	3181	6	ABN59646 Novel hum
5	1995	74.5	1995	5	AAF89709 Nucleotid
6	1294.2	48.3	2259	4	AAI88903 Human pol
7	945	35.3	1311	7	AAH188903 Human CDN
8	945	35.3	1311	7	ABX73184 Human nov
9	876.2	32.7	2186	7	AAH52592 Human cod
10	727	27.1	816	4	AAH06828 Human CDN
11	583	21.8	824	4	AAH06846 Human CDN
12	536.6	20.0	693	4	AAH26304 Human CDN
13	536.6	20.0	693	7	ABX73645 Human nov
14	495.8	18.5	575	3	AAH44336 Human sec
15	488.8	18.2	518	4	AAH11859 Human CDN
16	450	16.8	825	4	AAH58374 Human imm
17	394.4	14.7	476	8	ACH14968 Human adu
18	310.2	11.6	357	7	AAH26143 Human CDN
19	310.2	11.6	357	7	ABX73484 Human nov
20	273.6	10.2	449	8	ACH26450 Human adu
21	256.4	9.6	354	4	AAH26563 Human CDN
22	256.4	9.6	354	4	AAH26144 Human CDN
23	256.4	9.6	354	7	ABX73904 Human nov

24	256.4	9.6	354	7	ABX73485 Human nov
25	170	6.3	13836	5	ABH16633 Human ner
26	82.2	3.1	617	6	ABN60913 Human can
27	60	2.2	60	6	ABN33472 Human spl
28	52.4	2.0	2000	7	ADA71938 Rice gene
29	51.6	1.9	2000	7	ADA71938 Rice gene
30	49.4	1.8	612	4	AAH71471 Human cer
31	49	1.8	434	7	ABX38360 Bovine ES
32	48.2	1.8	7455	6	ABL33759 Human imm
33	47.8	1.8	309	5	ABV44994 Human pro
34	47.8	1.8	312	7	ABX43249 Bovine ES
35	47.8	1.8	474	5	ABV56492 Human pro
36	47.8	1.8	556	5	ABV40163 Human pro
37	47.8	1.8	556	5	ABV40063 Human pro
38	47.8	1.8	556	5	ABV42105 Human pro
39	47.8	1.8	556	5	ABV43601 Human pro
40	47.2	1.8	424	7	ABX46053 Bovine ES
41	47.2	1.8	612	4	AAH71471 Human cer
42	47.2	1.8	2122	7	ACC79105 Human sec
43	47.2	1.8	6215	6	ABL33190 Human cdn
44	47	1.8	2497	3	AAH61197 Human imm
45	46.6	1.7	1290	8	AAH57083 Human aid

ALIGNMENTS

RESULT 1
AAF30352
ID AAF30352 standard; cDNA; 2679 BP.
XX AAF30352;
AC AAF30352;
XX
DT 14-MAY-2001 (first entry)
XX
DE Human chfr cDNA encoding checkpoint with FHA and ring finger protein.
XX
KW Checkpoint with forkhead associated domain and ring finger; Chfr; human; mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening;
KW ubiquitin-protein ligase; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 91..2085
FT mutation /*tag= a
FT /*tag= b
FT /*note= "alters codon GTG (Val) to ATG (Met) in chfr: from human cancer cell line U2OS"
XX
XX WO200109150-A2.
XX
XX 08-FEB-2001.
XX
XX 14-JUN-2000; 2000WO-US016391.
XX
XX 29-JUL-1999; 99US-0146194P.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Halazonetis T, Scolnick D;
XX WPI; 2001-182927/18.
XX P-PSDB; AAB20219.
XX
XX Novel nucleic acid sequence of mitotic checkpoint gene encoding a checkpoint with forkhead-associated domain and ring finger protein, for diagnosing tumorigenic cells and in screening for anticancer drugs.
XX Claim 2(a); Fig 4A-D; 85pp; English.
XX
XX The present sequence is that of cDNA encoding the human mitotic

QY	2122	GTCAAGAGAGCGTGTGTTTTTAAATAACAGAGACAAGCACGTCAGGTGTTTTTACAGCCCC	2181
Db	1089	GTCAAGAGAGCGTGTGTTTTTAAATAACAGAGACAAGCACGTCAGGTGTTTTTACAGCCCC	1148
QY	2182	CTCAGGGAAAGGACCGCAGGGTCTCCGACAGGTGCTCTGGGGTGACTCTCTCTGGAGCTT	2241
Db	1149	CTCAGGGAAAGGACCGCAGGGTCTCCGACAGGTGCTCTGGGGTGACTCTCTCTGGAGCTT	1208
QY	2242	TTTACCCCTCTGAGTGAGACCCCTCCGAGAGCCCGGGGGCCGAGCCGCCCTCTGGTG	2301
Db	1209	TTTACCCCTCTGAGTGAGACCCCTCCGAGAGCCCGGGGGCCGAGCCGCCCTCTGGTG	1268
QY	2302	AGCGCTGGGACAGGGCTCTGTTGGCATCAGCAGCAGAGAGCGAAGCCTTTCTGTAACATGC	2361
Db	1269	AGCGCTGGGACAGGGCTCTGTTGGCATCAGCAGCAGAGAGCGAAGCCTTTCTGTAACATGC	1328
QY	2362	GGCCGTCGCCCGAGAGGGGCACTTTGCTCTTTTGTACATTTCCGAAACTACAGTTAA	2421
Db	1329	GGCCGTCGCCCGAGAGGGGCACTTTGCTCTTTTGTACATTTCCGAAACTACAGTTAA	1388
QY	2422	AGCAGAACTCTGTTTTTCAAGAAAGTTTCAAGGAGAGAGGGCAAGTTTATCAAAAACATT	2481
Db	1389	AGCAGAACTCTGTTTTTCAAGAAAGTTTCAAGGAGAGAGGGCAAGTTTATCAAAAACATT	1448
QY	2482	GTTTCAGGAGAGGGAGCATTAAGTTTACAGCCTACAGGAGGTACACAATATCTGCTGCT	2541
Db	1449	GTTTCAGGAGAGGGAGCATTAAGTTTACAGCCTACAGGAGGTACACAATATCTGCTGCT	1508
QY	2542	GGGAAACCCACAGCATTTTATCTATTTTATTTTAAATAGGTTGGTGCTTATCTTCTAA	2601
Db	1509	GGGAAACCCACAGCATTTTATCTATTTTATTTTAAATAGGTTGGTGCTTATCTTCTAA	1568
QY	2602	TAAGATTTAAATGTCACAACTGTAGCACAATAATATTAATTTTACAAATTGAC	2661
Db	1569	TAAGATTTAAATGTCACAACTGTAGCACAATAATATTAATTTTACAAATTGAC	1628
QY	2662	AAAAA 2666	
Db	1629	TAAAA 1633	

Search completed: May 14, 2004, 17:19:06
Job time : 16450 secs

```
QY 1714 GGTGCAAGTGTCTGGACGGGTGCTGAACAAACAGCTAGAGTCAGACATCTGAAG 1773
Db 4662 GGTGCAAGTGTCTGGACGGGTGCTGAACAAACAGCTAGAGTCAGACATCTGAAG 4721
QY 1774 AATTACCTGGCAACAGAGGTTTGACATGGAAAAACATGTTGACCGAGAGCTCGTGCT 1833
Db 4722 AATTACCTGGCAACAGAGGTTTGACATGGAAAAACATGTTGACCGAGAGCTCATGCT 4781
QY 1834 CTCACGCGGGAGTGTCTCTGCTGCTGATACAGAGTCAGGAGACACCGTCTCTGCT 1893
Db 4782 CTCACGCGGGAGTGTCTCTGCTGCTGATACAGAGTCAGGAGACACCGTCTCTGCT 4841
QY 1894 TACTCTCTGGCTCGCAGCTTCGGTGAAGTCTATCATAGTATCGATATCGGAGACATTCCT 1953
Db 4842 TACTCTCTGGCTCGCAGCTTCGGTGAAGTCTATCATAGTATCGGAGACATTCCT 4901
QY 1954 GCTTCGAGTTCGCGAGTGGCGGTATCATCCGCTCTGCTGCTGCTGCTGCTGCTGCT 2013
Db 4902 GCTTCGAGTTCGCGAGTGGCGGTATCATCCGCTCTGCTGCTGCTGCTGCTGCTGCT 4961
QY 2014 CCGACTCAGGTGAAAGCTCACACGCCATGAATTCATATCATATCTGTGAACAGACAAG 2073
Db 4962 CCGACTCAGGTGAAAGCTCACACGCCATGAATTCATATCATATCTGTGAACAGACAAG 5021
QY 2074 TTCAAAACTAAGCATCCAGAGCCCTGAGGAGCTTTGAGCACTGAGAGTGAAGAGCG 2133
Db 5022 TTCAAAACTAAGCATCCAGAGCCCTGAGGAGCTTTGAGCACTGAGAGTGAAGAGCG 5081
QY 2134 TGTTTTTAAATACAGAGACAGCAGTCAAGGTGTTTTCACAGCCCTCGAGGAGAGG 2193
Db 5082 TGTTTTTAAATACAGAGACAGCAGTCAAGGTGTTTTCACAGCCCTCGAGGAGAGG 5141
QY 2194 AGCAGGCTCTCCGACAGTGTCTGGGGTGACTCTTCTGTGGAGCTTTTACCTCTGA 2253
Db 5142 AGCAGGCTCTCCGACAGTGTCTGGGGTGACTCTTCTGTGGAGC--TTTACCTCTGA 5199
QY 2254 GTGAGACCTCCCGAGAGCCCGGGGGCGGAGCCGCTCTGCTGAGCGCTGGGCGAG 2313
Db 5200 GTGAGACCTCCCGAGAGCCCGGGGGCGGAGCCGCTCTGCTGAGCGCTGGGCGAG 5259
QY 2314 GGCTCGTGGTGCATCAGCAGCAGAGCAAGCCCTTTCTGTAACATGGCGCCGCTCCGCC 2373
Db 5260 GGCTCGTGGTGCATCAGCAGCAGAGCAAGCCCTTTCTGTAACATGGCGCCGCTCCGCC 5319
QY 2374 GAGAGGGCAGTTTGTCTTTTGTACATTTTCCGAACACTACAGTTAAAGCAGAGTCTG 2433
Db 5320 GAGAGGGCAGTTTGTCTTTTGTACATTTTCCGAACACTACAGTTAAAGCAGAGTCTG 5379
QY 2434 TTTTCAGGAAAGTTTCAAGGAGAGGCGCAAGTTTATCAAAAAACATTTGTTTCAGGAGAA 2493
Db 5380 TTTTCAGGAAAGTTTCAAGGAGAGGCGCAAGTTTATCAAAAAACATTTGTTTCAGGAGAA 5439
QY 2494 GGGAGCATAGTTTACAGCTACAGGAGTACAGATATCTGCTGCTGGGAAACCCACA 2553
Db 5440 GGGAGCATAGTTTACAGCTACAGGAGTACAGATATCTGCTGCTGGGAAACCCACA 5499
QY 2554 GCATTTTATCTATTTTATTTTATAGTTTGTGCTTATCTTCTTAATAAGATTAAAT 2613
Db 5500 GCATTTTATCTATTTTATTTTATAGTTTGTGCTTATCTTCTTAATAAGATTAAAT 5559
QY 2614 GTCACAACTGAGCAGCAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2666
Db 5560 GTCACAACTGAGCAGCAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5612
```

RESULT 15

```
AK090948
LOCUS AK090948 2257 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ33629 fis, clone BRAMY2022454, highly similar
to Homo sapiens cell cycle checkpoint protein CHEF mRNA.
ACCESSION AK090948
VERSION AK090948.1 GI:21749208
KEYWORDS oligo capping; fis (full insert sequence).
```

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
Kodaira,H., Furiya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshina,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
TITLE NEDO human cDNA sequencing project
JOURNAL 2 (bases 1 to 2257)
REFERENCE Isogai,T. and Yamamoto,J.
AUTHORS Direct Submission
TITLE Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
JOURNAL Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(Is-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1. 2257
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAMY2022454"
/tissue_type="amygdala"
/clone_lib="BRAMY2"
/note="Cloning vector: pME18SFL3"
ORIGIN
Query Match 35.9%; Score 961.8; DB 9; Length 2257;
Best Local Similarity 99.8%; Pred. No. 4.1e-210;
Matches 963; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1702 GAGCTCAACCTGGTGACAGTGTCTGGACGGGTGCTGACACACACAGCTACGAGTCA 1761
Db 669 GAGCTCAACCTGGTGACAGTGTCTGGACGGGTGCTGACACACACAGCTACGAGTCA 728
QY 1762 GACATCTGAAGAAATACCTGGCAACAGAGGTTTGACATGGAAAAACATGTTGACCGAG 1821
Db 729 GACATCTGAAGAAATACCTGGCAACAGAGGTTTGACATGGAAAAACATGTTGACCGAG 788
QY 1822 AGCCTCGTGCTCTCCAGCGGGGAGTGTCTGCTGCTGATTAACAGATCAGCGAGAC 1881
Db 789 AGCCTCGTGCTCTCCAGCGGGGAGTGTCTGCTGCTGATTAACAGATCAGCGAGAC 848
QY 1882 ACCGTTCTGTGTACTGCTGTGGCCCTCCGAGCTTCGCTGAGCTGACCTATCAGTATCG 1941
Db 849 ACCGTTCTGTGTACTGCTGTGGCCCTCCGAGCTTCGCTGAGCTGACCTATCAGTATCG 908
QY 1942 CAGAACATTCCTGCTTCCGAGTTCGAGTGGCGGTAAACATCCCGTCTGCTGCTGCTACTCG 2001
Db 909 CAGAACATTCCTGCTTCCGAGTTCGAGTGGCGGTAAACATCCCGTCTGCTGCTGCTACTCG 968
QY 2002 GCGCGTAATCGCGCACTCAGGTGAAGCTCACCAGCCATGAATTCATATATCTCT 2061
Db 969 GCGCGTAATCGCGCACTCAGGTGAAGCTCACCAGCCATGAATTCATATATCTCT 1028
QY 2062 GAACAGACAGAGTTCAAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTTTCAGCAGTGGAG 2121
Db 1029 GAACAGACAGAGTTCAAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTTTCAGCAGTGGAG 1088
```

Db	1447	GGCACCTTCAACATCTGCCAGTTGCCAAACAGACAGCCAGATTACATGTGCCCTCTTCA	1506
Qy	1527	AGNAGACCGCCTGTGCACTGCTCTTCCAGCCCATGCCCGACGGAGACGGAGCG	1586
Db	1507	AGGAAGCCATGCCATATGCACTGCTGCTTCCAGCCTATCCCTGACCGGAGACTGAACG	1566
Qy	1587	CGACAGGACCCCGGTGTGCGCCCTCAGCAGTGTGCGGTGTGCTGCTGAGCCTTTCTGCCA	1646
Db	1567	GGACAGGATCCCGCGGTGCGCCCTCAGCAGTGTGCGGTGTGCTGCTGAGCCTTTCTGCCA	1626
Qy	1647	CTGTATCTGGGGCTGCAACCGGACCGGTGCTAGCGCTGCTGGCCCCCGTTTGTGAGCT	1706
Db	1627	CCTGTATCTGGGGCTGCAAGGAGCTGCTGCTTTGGCTGTGTGCCCATCTCTGTGAGCT	1686
Qy	1707	CAACCTGGGTGCAAAAGTGTCTGGACCGCGTGTGAAACAACAAGCTTACAGTCAGACAT	1766
Db	1687	CAACCTGGGGACAAGTGTCTGGATGAGTGTGAAACAATPACAAATATGAATCGGACAT	1746
Qy	1767	CCTGAAGAAATTAATCTGGCAACAGAGGTTTGAATGGAAGAAAATGTTGTACCGAGAGCCT	1826
Db	1747	CCTGAAGAAATTAATCTGGCAACAGAGGTTTGAATGGAAGAAAATGTTGTGACAGAGTCT	1806
Qy	1827	CGTGGCTCTCCAGCGGGAGTGTCTTCTGCTGTCTGATTACAGAGTCACGGGACACCGT	1886
Db	1807	CCTGGCTCTGACGACAGGTGTATTATGCTGTCTGATTAAGAAATCACTGGAAATCTGT	1866
Qy	1887	TCTGTGTTACTGTGTGGCTGCGCAGCTTCCGFGAGCTGACCTATCAGTATCGGCAGAA	1946
Db	1867	GCTGTGTTACTGTGTGGCTGCGTAGCTTCCGAGAGCTGACCTACCAGTATCGTCAGAA	1926
Qy	1947	CATTCTGCTTCCAGTGTCCAGTGGCGGTAAACATCCGCTCTGACTGCTACTGGGGCGG	2006
Db	1927	CATTCTGCTTCTGAGTGTCCAGTGTACTGTAAACATCCGCTCTGACTGCTACTGGGGCGG	1986
Qy	2007	TAACTGCCGCACTCAGGTGAAGCTCACAGCCCATGAATTCATCATATCTGTGAACA	2066
Db	1987	CAACTGTCGCACTCAGGTGAAGCTCACAGCCCATGAATTCATCATATCTGTGAACA	2046
Qy	2067	GACAGGTTCCAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTCACACCTGAGAGTGAA	2126
Db	2047	AACAGGTTTCAGAACTGAAATTTGGAGAGACAGCTGAAACAGTTACAGACACAGGAG	2106
Qy	2127	GAGAGCGTGTTTTTAAAATACAGAGACAGCAGCTCAAGGTGTTTTACAGCCCCCTGAG	2186
Db	2107	GCAAT-GTTTTTTTAAAGGCCCCATACAAACACTTGAAGGGCATTTTACAGTTTCCCT	2165
Qy	2187	GGAAAGGACGACGGTCTCCAGACAGTGTCTTGGGTGACTCTCTCTGTGAGCTTTTAC	2246
Db	2166	ACGACAGCACTGGGGGTTCTGCTGACGCACTGGGGTGTGGAATTTTCCCTTTGTGAGCATC	2225
Qy	2247	CCTCTGAGTGAGACCTCCACAGAGCCCCGGGGCGCCAGCCCCCTCTGTGTGAGCGC	2306
Db	2226	ATCCTCAGTGACACACACACACCTGCCCCAGTAAAC-CTGGAGGCGCTCTGGCCACACC	2284
Qy	2307	TGGCGAGGGCTGTGTGGCATCAGCAGCAGAG-----ACGAAGCC	2347
Db	2285	AGCCTTCTAGTCGGTGTAAAGAGGATCAGTAGACTGAGTTCTCCCTCCCTCCCACTGGACC	2344
Qy	2348	TTTCTGTAAACATGGCGCGTCCCGCCGAGAGGGCAGTTTGTCTCTTTTGTACATTTTC	2407
Db	2345	GCTCTCCCTCCTGTGCTGTAGGCAGGTGACAGTGCAGGTTTGTACTTTTGTACTTTT--	2402
Qy	2408	GAATCTACAGTAAAGCAGAAAGTCTGTTTTCAGAAAGTGTTCAGGGAGAGGGCAAGT	2467
Db	2403	----CTACCTTGTAACCTGACTTTTGGTTTAAATGGGAAGCTCATAGGAAAGTTGTGAGC	2458
Qy	2468	TTATCAAAAACATGTTTTCAGGAG--AAGGAGCATAGTTTACAGCTACAGACGT-A	2524
Db	2459	TGTTCTCAGACGCATCTTCCGGAAGGAAGGAGCATACGTTTACAGCTCCAGGGGTACA	2518
Qy	2525	CACAAATCTCTGCTGCTGGGAAAAACACAGCATTTTATCTATTTTTTTATTTTAAATAGGTT	2584

D	b		2519	TAGGATAGTCTCCTGTCTGGAAAGACACAGCCCTTTTGTAATAATTTTTTATTTAATAAGTT	2576
Q	y		2585	TGGTGCTTATCTCTCAATAAGATTTAAATGTCAAACTGTAGCACAAATAATAATTT	2644
D	b		2579	TGGTGCTTATCTCTCAATAAGATTTAAATATTAAAACTGTAGCACAAAATACTATAATTT	2638
Q	y		2645	ATAATTTCACAANTGCARAAAAAAAAAAAAA	2675
D	b		2639	ATAATTTACAAATTGACTAAATTAGGAATA	2669
 RESULT 14					
BD183291					
LOCUS BD183291 6235 bp DNA linear PAT 17-JUN-2003					
DEFINITION Novel genes and proteins encoded by the genes.					
ACCESSION BD183291					
VERSION BD183291.1 GI:31875491					
KEYWORDS JP 2002345492-A/4.					
SOURCE Homo sapiens (human)					
ORGANISM Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
1 (bases 1 to 6235)					
REFERENCE Chara,O., Nagase,T. and Nakajima,D.					
AUTHORS Novel genes and proteins encoded by the genes					
TITLE Patent: JP 2002345492-A 4 03-DEC-2002;					
JOURNAL KAZUSA DNA RESEARCH INSTITUTE					
COMMENT OS Homo sapiens (human)					
PN JP 2002345492-A/4					
PD 03-DEC-2002					
PF 26-FEB-2002 JP 2002049009					
PI OSAKI CHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA					
PC C12N15/09,C07K14/47/A61K31/711,A61K38/00,A61K48/00,A61P25/00,					
PC A61P25/14,					
PC A61P25/18,A61P35/00,C12N15/00,A61K37/02					
CC Novel genes and proteins encoded by the genes FH Key					
Location/Qualifiers					
FT CDS Location/Qualifiers (50), (1414).					
FEATURES					
source					
1..6235					
/organism="Homo sapiens"					
/mol_type="genomic DNA"					
/db_xref="taxon:9606"					
 ORIGIN					
Query Match 47.8%; Score 1279.8; DB 6; Length 6235;					
Best Local Similarity 99.5%; Pred. No. 3.2e-283;					
Matches 1306; Conservative 0; Mismatches 2; Indels 5; Gaps 2;					
Q	y		1354	AGCCAGCATACGTCGTGTGCCGACGTCTCTGAGTACAGAAGCGCGCGCAGCCT	1413
D	b		4305	AGCCAGCCATACGTCGTGTGCCGACGTCTCTGAGTACAGAAGCGCGCGCAGCCT	4364
Q	y		1414	CCCACCTCCAGCACCCGAGGGCGACGAGCCCCACAGGCCCTTGGGGGATGCACC	1473
D	b		4365	CCCACCTCCAGCACCCGAGGGCGACGAGCCCCACAGGCCCTTGGGGGATGCACC	4424
Q	y		1474	TCCACGTCCTGACGCTGACGACAGCAGTCCAGGATTCGTGTGCCCTCTGCAAGGAAGC	1533
D	b		4425	TCCACGTCCTGACGCTGACGACGAGTTCGTGTGCCCTCTGCAAGGAAGC	4481
Q	y		1534	CACGCCCTGTGTGCACTCTGTGTTCCAGCCATGCCGAGCGGAGCGGAGCGGAGCAG	1593
D	b		4482	CACGCCCTGTGTGCACTCTGTGTTCCAGCCATGCCGAGCGGAGCGGAGCGGAGCAG	4541
Q	y		1594	GACCGCGGTGTGCGCCCTCAGCAGTGTGCGGTCTGCTGCAGCCTTTCTGCCACCTGTAC	1653
D	b		4542	GACCGCGGTGTGCGCCCTCAGCAGTGTGCGGTCTGCTGCAGCCTTTCTGCCACCTGTAC	4601
Q	y		1654	TGGGGCTGCACCCCGGACCGGCTGTACGGCTGTCTGCCCCCGTTTGTAGCTCAACCTG	1713
D	b		4602	TGGGGCTGCACCCCGGACCGGCTGTACGGCTGTCTGCCCCCGTTTGTAGCTCAACCTG	4661

QY		2665 AA 2666	
Db		1501 AA 1502	
RESULT 13			
BC049792			
LOCUS		3194 bp mRNA linear ROD 07-OCT-2003	
DEFINITION		Mus musculus RIKEN cDNA 573048M20 gene, mRNA (cdna clone MGC:59419 IMAGE:6510480), complete cds.	
ACCESSION		BC049792	
VERSION		BC049792.1	
KEYWORDS		G1:29437110	
SOURCE		MGC.	
ORGANISM		Mus musculus (house mouse)	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		1 (bases 1 to 3194) Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zebins,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Capetonek,M., Soares,M.B., Donaldson,M.P., Casavant,T.L., Sametz,T.E., Brownstein,M.J., Usdin,T.B., Loquellano,N.A., Peters,G.J., Carrinci,P., Prange,C., Raha,S.S., Bosak,S.A., McEwan,P.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McGowan,K.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalls,D.E., Scherach,A., Schein,J.E., Jones,S.J. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
TITLE		human and mouse cDNA sequences	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
MEDLINE		22388257	
PUBMED		12477932	
REFERENCE		2 (bases 1 to 3194)	
AUTHORS		Straussberg,R.	
TITLE		Direct Submission	
JOURNAL		Submitted (31-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK		NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT		Contact: MGC help desk Email: cgabbs-r@mail.nih.gov Tissue Procurement: Dr. David Rowe CDNA Library Preparation: Invitrogen Corp CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Rita Kutache, Oliver Lee, Soo Len Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven Ness, Pawan Pandon, Anna-Lisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAX Plate: 109 Row: 1 Column: 17 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27370049, Location/Qualifiers	
FEATRES			

RESULT 12	HSM802304	2297 bp	mrna	linear	PRI 18-FEB-2000
LOCUS	HSM802304				
DEFINITION	Homo sapiens mRNA; cDNA DKFZp43N2420 (from clone DKFZp43N2420) ; partial cds.				
ACCESSION	AL137561				
VERSION	AL137561.1	GI:6808265			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 2297)				
AUTHORS	Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and Wiemann, S.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152				

Query Match	56.0%	Score 1500.4	DB 9	Length 2297
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1501	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1165	GTGGAGCATACTCTATCCAGCATCCAGACACAGATCGCGAGTCAAGAAGATGTGCAAAAGT	1224	
DB	1	GTGGAGCATACTCTATCCAGCATCCAGACACAGATCGCGAGTCAAGAAGATGTGCAAAAGT	60	
QY	1225	ATGGATGCCAGGAATAAAATCACTCAAGACATGCTCCAGCCCCAAAGTCAAGCGGTCTTTT	1384	
DB	61	ATGGATGCCAGGAATAAAATCACTCAAGACATGCTCCAGCCCCAAAGTCAAGCGGTCTTTT	120	
QY	1285	TCTGATGAGAGGGAGTTTCAGAGGACCTGCTGGAGCTGTTCAGACGTTGACAGTGAGTCC	1344	
DB	121	TCTGATGAGAGGGAGTTTCAGAGGACCTGCTGGAGCTGTTCAGACGTTGACAGTGAGTCC	180	
QY	1345	TCAGACATTAGCCAGCCATACGTCGTGTGCGCGAGTGTCTGAGTACAGAAAGCAGCGC	1404	
DB	181	TCAGACATTAGCCAGCCATACGTCGTGTGCGCGAGTGTCTGAGTACAGAAAGCAGCGC	240	
QY	1405	GCGCAGCTTCCCACTGCCCCAGACCCCGAGGGCGAGCCAGGAGCCCCACAGGCCCTTGGGG	1464	
DB	241	GCGCAGCTTCCCACTGCCCCAGACCCCGAGGGCGAGCCAGGAGCCCCACAGGCCCTTGGGG	300	
QY	1465	GATGCAACCTTCCACGTGCCGTTCAGCCTTCAGCAGCAGTCCAGGATTACGTGTGCCCTCTG	1524	
DB	301	GATGCAACCTTCCACGTGCCGTTCAGCCTTCAGCAGCAGTCCAGGATTACGTGTGCCCTCTG	360	
QY	1525	CAAGGAAGCCAGCCCTGTGACCTGCTGTTCCAGCCCATGCCCGACCGAGAGCGGAG	1584	
DB	361	CAAGGAAGCCAGCCCTGTGACCTGCTGTTCCAGCCCATGCCCGACCGAGAGCGGAG	420	

Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and
Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2448)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 252-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

source

1..2448
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST12033505"
/tissue_type="testis"
/clone_lib="TEST12"
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 71.6%; Score 1919.4; DB 9; Length 2448;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2227; Conservative 0; Mismatches 6; Indels 284; Gaps 2;
218 GACGAGGTGGACCTTTCCTTCCCGAGCAATAAATGCTCTCTGGAGATCACTGTAGAA 277
83 GACCTGTTGGACCTTTCCTTCCCGAGCAATAAATGCTCTCTGGAGATCACTGTAGAA 142
278 TTGTAGTGGATGAAAAATCAGGTGAGTGACATGGAAGATACAGACACAGTGGACAG 337
143 TTGTAGTGGATGAAAAATCAGGTGAGTGACATGGAAGATACAGACACAGTGGACAG 202
338 TGATTAACAGCTCAAGTGTGTTAAGAGCAGACATGCCCTTTACAGACTGGGGATGTCA 397
203 TGATTAACAGCTCAAGTGTGTTAAGAGCAGACATGCCCTTTACAGACTGGGGATGTCA 262
398 TCTACTTGGTGTACAGGAAGATGAACCGGAACAACAACGTGGGATACCTCTATGAATCTT 457
263 TCTACTTGGTGTACAGGAAGATGAACCGGAACAACAACGTGGGATACCTCTATGAATCTT 322
458 TAAGTGAACAGCAGGATGACACAGAAATCCTTTTGAAGCTAACAAGAAATGTGTTC 517
323 TAAGTGAACAGCAGGATGACACAGAAATCCTTTTGAAGCTAACAAGAAATGTGTTC 356
518 ATGGGACCAAGATACCTCAGGTGAGTGAGGGGCGAGGGGCGATCCCGGGTCCCTC 577
357 ----- 356
578 CGTGTGCCCGCCACTCAGGTGTGCTTTGAGGAACACAGACCATCAACATCGAGCTCAG 637
357 ----- 356
638 ACCTCTCCCGACAGCCTCGGCTCTTCCACGAGACCTTTCTCTGCGAGGCGAGAGGTT 697
357 ----- 356
698 CCTCAGTTGGGTCTGGGGTGGTGCACTCCCTTAAGGAAGTGGTCCCTCTGTGG 757
357 -----TGGGTCTGGGGTGGTGCACTCCCTTAAGGAAGTGGTCCCTCTGTGG 406
758 CAAGTGAAGTCTCCAGCTTTGCTCAGTCTCTCCAGACAGAAAGACTGCGTCTCTTT 817
407 CAAGTGAAGTCTCCAGCTTTGCTCAGTCTCTCCAGACAGAAAGACTGCGTCTCTTT 466

467 CGTGGTGAACCCAGGATCAGGAGGATTTGGAGCCGTTGAAGAAATGAGAGGAG 526
878 ATGGGGACCTTGACCTGAACGGGCAAGTTGTTGGTCGACACACCGCTAGAAATGCCCAA 937
527 ATGGGGACCTTGACCTGAACGGGCAAGTTGTTGGTCGACACACCGCTAGAAATGCCCAA 586
938 CCGTCCAGAGAGCTCAGAGCAGCGGCTGGGAAAGCCAGACAAAGATGAGAGAGCGCTGA 997
587 CCGTCCAGAGAGCTCAGAGCAGCGGCTGGGAAAGCCAGACAAAGATGAGAGAGCGCTGA 646
998 CATGCATCATCTGCAGAGACCTGCTGCAAGCTGGGAGTGGTTGAGAGCCCTGCATGACA 1057
647 CATGCATCATCTGCAGAGACCTGCTGCAAGCTGGGAGTGGTTGAGAGCCCTGCATGACA 706
1058 CGTTCTGGCGGCTTGCTTACTCGGCTGATGAGAGCGCTGCTCCTGCTGCTTACCTGCC 1117
707 CGTTCTGGCGGCTTGCTTACTCGGCTGATGAGAGCGCTGCTCCTGCTGCTTACCTGCC 766
1118 GCTGTCCGCTGGAGCGGATCTGTAAACCAATCCTCAACAACTCGTGGAAAGATACC 1177
767 GCTGTCCGCTGGAGCGGATCTGTAAACCAATCCTCAACAACTCGTGGAAAGATACC 826
1178 TCATCCAGATCCAGACAGAGTGCAGTGAAGATGTCGAAGATGATGATGATGCCAGGA 1237
827 TCATCCAGATCCAGACAGAGTGCAGTGAAGATGTCGAAGATGATGATGATGCCAGGA 886
1238 ATAAATCACTCAAGACATGCTGACGCCCAAGTCAAGCGGCTCTTTTCTGATGAAGAAG 1297
887 ATAAATCACTCAAGACATGCTGACGCCCAAGTCAAGCGGCTCTTTTCTGATGAAGAAG 946
1298 GGAGTTCAGAGACCTGCTGAGAGCTGTGACAGTGTGACAGTGTGACAGATGATGCC 1357
947 GGAGTTCAGAGACCTGCTGAGAGCTGTGACAGTGTGACAGTGTGACAGATGATGCC 1006
1358 AGCCATAGCTGTGTCGGGAGTGTCTGAGTACAGAAAGCAGCGCGGCGAGCTCCCC 1417
1007 AGCCATAGCTGTGTCGGGAGTGTCTGAGTACAGAAAGCAGCGCGGCGAGCTCCCC 1066
1418 ACTGCCAGACCCAGAGGCGGAGCCAGAGGCCCAAGGCCCCAGGCCCCGAGGATGCCCTCCA 1477
1067 ACTGCCAGACCCAGAGGCGGAGCCAGAGGCCCAAGGCCCCAGGCCCCGAGGATGCCCTCCA 1126
1478 CGTCCGCTCAGCTCAGACAGCAGTCCAGAGTACGTTGTCCTCTGCAAGGAGCCACG 1537
1127 CGTCCGCTCAGCTCAGACAGCAGTCCAGAGTACGTTGTCCTCTGCAAGGAGCCACG 1186
1538 CCCTGTGACCTGTGCTTCCAGGCCATGCCCGACCCGAGAGCGGAGCGGAGCAGGACC 1597
1187 CCCTGTGACCTGTGCTTCCAGGCCATGCCCGACCCGAGAGTGGAGCGGAGCAGGACC 1246
1598 CGCGTGTGCCCTCAGCAGTGTGCGGCTGCGCTGACGCTTCTGCGACCTGTACTGGG 1657
1247 CGCGTGTGCCCTCAGCAGTGTGCGGCTGCGCTGACGCTTCTGCGACCTGTACTGGG 1306
1658 GCTGCACCCGAGCCGCTGCTACGCGCTGCTGCGCCCGCTTTTGT----- 1701
1307 GCTGCACCCGAGCCGCTGCTACGCGCTGCTGCGCCCGCTTTTGTGTGTGTCGCGGAGCAGT 1366
1702 -----GAGCTCAA 1709
1367 GCTTTCTCGCAGCATATTATTAACAGGTAAATGCCCCCTTCATCTCGGAGAGCTCAA 1426
1710 CTTGGGTGACAGTGTCTGGAGCGGCTGTAACCAACAGCTACGAGTACGAGCATCCT 1769
1427 CTTGGGTGACAGTGTCTGGAGCGGCTGTAACCAACAGCTACGAGTACGAGCATCCT 1486
1770 GAAGAATTACCTGGCAACAGAGTTTGACATGGAAACATGTTGACCGAGCGCTCGT 1829
1487 GAAGAATTACCTGGCAACAGAGTTTGACATGGAAACATGTTGACCGAGCGCTCGT 1546
1830 GCGCTCTCCAGCGGAGGTTTCTGCTGCTGATTAACAGAGTCCAGGAGACACCGTCTCT 1889
1547 GCGCTCTCCAGCGGAGGTTTCTGCTGCTGATTAACAGAGTCCAGGAGACACCGTCTCT 1606

Qy 1890 GTGTTACTGCTGTGGCCTGCGCAGCTTCGTTGAGCTGACCTATCAGTATCGGCAGAACAT 1949
Db 1607 GTGTTACTGCTGTGGCCTGCGCAGCTTCGTTGAGCTGACCTATCAGTATCGGCAGAACAT 1666
Qy 1950 TCGTGTTCGAGTTGCGAGTGGCGGTAAACATCCCGTCTGACTGCTACTTGGGGCGGTAA 2009
Db 1667 TCGTGTTCGAGTTGCGAGTGGCGGTAAACATCCCGTCTGACTGCTACTTGGGGCGGTAA 1726
Qy 2010 CTGCGCCACTCAGGTGAAAGCTCACCCAGCCATGAAATCAATCATATCTGTGAAACAGAC 2069
Db 1727 CTGCGCCACTCAGGTGAAAGCTCACCCAGCCATGAAATCAATCATATCTGTGAAACAGAC 1786
Qy 2070 AAGGTTCAAAAATAAGCATCAGAGGCGCTGAGAGCTTTTTCAGCACTGGAGGTGAAGAG 2129
Db 1787 AAGGTTCAAAAATAAGCATCAGAGGCGCTGAGAGCTTTTTCAGCACTGGAGGTGAAGAG 1846
Qy 2130 AGCGTGTGTTTAAATACAGAGCAGCAGCTCAAGTGTGTTTTCAGCGCCCTCGAGGGA 2189
Db 1847 AGCGTGTGTTTAAATACAGAGCAGCAGCTCAAGTGTGTTTTCAGCGCCCTCGAGGGA 1906
Qy 2190 AGGAGCGCAGGCTCTCCGACAGGTGCTCTGGGCTGACTCTTCTGTGGAGCTTTTACCCT 2249
Db 1907 AGGAGCGCAGGCTCTCCGACAGGTGCTCTGGGCTGACTCTTCTGTGGAGCTTTTACCCT 1966
Qy 2250 CTGAGTGAGACCTCTCCCGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2309
Db 1967 CTGAGTGAGACCTCTCCCGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2026
Qy 2310 GCAGGCTGCTGTGTGATCAGCAGCAGAGAGAGAGCTTTTCTGTAACATGCGGCGGCTCC 2369
Db 2027 GCAGGCTGCTGTGTGATCAGCAGCAGAGAGAGAGCTTTTCTGTAACATGCGGCGGCTCC 2086
Qy 2370 CGCGGAGAGGCGGCTGCTGCTGTTTGTACATTTTCCGAACTACAGTTAAAGCAGAAG 2429
Db 2087 CGCGGAGAGGCGGCTGCTGCTGTTTGTACATTTTCCGAACTACAGTTAAAGCAGAAG 2146
Qy 2430 TCTGTTTTCAGGAAAGTTTCAAGGAGAGAGGCAAGTTTATCAAAAACATGTTTTCAGG 2489
Db 2147 TCTGTTTTCAGGAAAGTTTCAAGGAGAGAGGCAAGTTTATCAAAAACATGTTTTCAGG 2206
Qy 2490 AGAAGGAGAGATTAAGTTTACAGCTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAG 2549
Db 2207 AGAAGGAGAGATTAAGTTTACAGCTACAGAGCAGAGAGAGAGAGAGAGAGAGAGAG 2266
Qy 2550 CACAGCATTTTACATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2609
Db 2267 CACAGCATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2326
Qy 2610 AAATGTCAAAAATGAGCAGCAAAATTAATAATTTATAATTTTACAAATTCAGCAAAA 2666
Db 2327 AAATGTCAAAAATGAGCAGCAAAATTAATAATTTATAATTTTACAAATTCAGCAAAA 2383

RESULT 11

AK097671

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AK097671 2448 bp mRNA linear PRI 15-JUL-2002
Homo sapiens cDNA FLJ40352 fis, clone TESTI2033505, highly similar
to Homo sapiens cell cycle checkpoint protein CHFR mRNA.

AK097671.1 GI:21757518
oligo capping; fis (full insert sequence).
Homo sapiens (human)

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.

1
Oshima, A., Takahashi-Fujii, A., Tanase, T., Inose, N., Takeuchi, K.,
Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hiro, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,

Qy	1636	CCTTTCTGCCACCTGTACTGGGCGTGCA	CCCGGACCGGCTGCTACGGCTGCCCTGGGCCCGG	1699
Db	1518	CCTTTCTGCCACCTGTACTGGGCGTGCA	CCCGGACCGGCTGCTACGGCTGCCCTGGGCCCGG	1577
Qy	1696	TTTTGTGAGCTCAACCTGGGCTGACAA	GCTGCTGGACCGGCTGTGAAACAAACAGCTAC	1755
Db	1578	TTTTGTGAGCTCAACCTGGGCTGACAA	GCTGCTGGACCGGCTGTGAAACAAACAGCTAC	1637
Qy	1756	GAGTCAGACATCCTGTGAAGAAATTAC	CTGGGCAACGAGGTTTGACATGGAAAAACATGTTG	1815
Db	1638	GAGTCAGACATCCTGTGAAGAAATTAC	CTGGGCAACGAGGTTTGACATGGAAAAACATGTTG	1697
Qy	1816	ACCGAGAGCCTCGTGGCTCTCCACGG	GGGAGTGTTCCTGCTGTCTGATTACAGAGTCACG	1875
Db	1698	ACCGAGAGCCTCGTGGCTCTCCACGG	GGGAGTGTTCCTGCTGTCTGATTACAGAGTCACG	1757
Qy	1876	GGAGACACCGTTCGTGTGTTACTGTCT	GGGCTGGGACGTTCCGTGAGCTGACCTATCAG	1935
Db	1758	GGAGACACCGTTCGTGTGTTACTGTCT	GGGCTGGGACGTTCCGTGAGCTGACCTATCAG	1817
Qy	1936	TATCGGACAGAACATTCCTGCTTCCG	AGTGTCCAGTGGCCGTAACATCCCGCTCGTACTGC	1995
Db	1818	TATCAGACAGAACATTCCTGCTTCCG	AGTGTCCAGTGGCCGTAACATCCCGCTCGTACTGC	1877
Qy	1996	TACTGGGGCCGTAACTCCGCGACCTC	AGGTGAAAGCTCACACGCCATGAATTCATCAT	2055
Db	1878	TACTGGGGCCGTAACTCCGCGACCTC	AGGTGAAAGCTCACACGCCATGAATTCATCAT	1937
Qy	2056	ATCTGTCAACAGACAAAGTTCAAAAA	CTTAAGCATCCAGAGCCCTGAGCAGCTTTTCAGCA	2115
Db	1938	ATCTGTCAACAGACAAAGTTCAAAAA	CTTAAGCATCCAGAGCCCTGAGCAGCTTTTCAGCA	1997
Qy	2116	CTGAGGTGAAGAGAGCGTGTTTTTAA	ATAATACAGACAGCAGCTCAAGGTGTTTTTCAC	2175
Db	1998	CTGAGGTGAAGAGAGCGTGTTTTTAA	ATAATACAGACAGCAGCTCAAGGTGTTTTTCAC	2057
Qy	2176	AGCCCCCTGAGGGAAGGACCGCAGG	CTCCGACAGGTGCTCTGGGGTGACTCTTCTGTG	2235
Db	2058	AGCCCCCTGAGGGAAGGACCGCAGG	CTCCGACAGGTGCTCTGGGGTGACTCTTCTGTG	2117
Qy	2236	GAGCTTTTATCCCTCTGAGTGAGACC	CTTCCCAGAGCCCGGGGCGCAGCCGCCCTC	2295
Db	2118	GAGC--TTTACCTCTCTGAGTGAGAC	CTTCCCAGAGCCCGGGGCGCAGCCGCCCTC	2175
Qy	2296	CTGCTGAGCGCTGGCGCAGGCTCGTG	TGGCATCAGCAGCAGAGCAAGAGCCCTTTCTGTA	2355
Db	2176	CTGCTGAGCGCTGGCGCAGGCTCGTG	TGGCATCAGCAGCAGAGCAAGAGCCCTTTCTGTA	2235
Qy	2356	ACATGGCGCGTCCGCGCAGAGGCGAG	TTTGCTCTTTTGTATCATTTTCCGAAACTAC	2415
Db	2236	ACATGGCGCGTCCGCGCAGAGGCGAG	TTTGCTCTTTTGTATCATTTTCCGAAACTAC	2295
Qy	2416	AGTTAAAGCGAAGTCTGTGTTTTTC	AGGAAAAAGTTTCAAGGGAGAGGCAAGTTATCAAA	2475
Db	2296	AGTTAAAGCGAAGTCTGTGTTTTTC	AGGAAAAAGTTTCAAGGGAGAGGCAAGTTATCAAA	2355
Qy	2476	AACATTGTTTCAGAGAAAGGAGCAT	TAAGTTTACAGGCTACAGGACGTACACATATCCT	2535
Db	2356	AACATTGTTTCAGAGAAAGGAGCAT	TAAGTTTACAGGCTACAGGACGTACACATATCCT	2415
Qy	2536	GCTGCTGGGAAAAACACAGCATTTT	TATCTATTTTTTATATGTTTGGTGCTTATC	2595
Db	2416	GCTGCTGGGAAAAACACAGCATTTT	TATCTATTTTTTATATGTTTGGTGCTTATC	2475
Qy	2596	TTCTTAATAAGATTTTAAATGTCA	CAAACTGTAGCAAAATAATAATTTATAATTTACAA	2655
Db	2476	TTCTTAATAAGATTTTAAATGTCA	CAAACTGTAGCAAAATAATAATTTATAATTTACAA	2535
Qy	2656	ATTGACAAAA	2666	
Db	2536	ATTGACTAAAA	2546	

RESULT 10
 AX834949
 LOCUS 2448 bp DNA linear PAT 15-DEC-2003
 DEFINITION Sequence 2073 from Patent EP1347046.
 AX834949
 ACCESSION
 VERSION AX834949.1 GI:39921084
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
 Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
 Tanemichi, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and
 Masuho, Y.
 TITLE Full-length cDNA sequences
 JOURNAL Patent: EP 1347046-A 2073 24-SEP-2003;
 Research Association for Biotechnology (JP)
 FEATURES
 source Location/Qualifiers
 1..2448
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN	Query Match	71.6%; Score 1919.4; DB 6; Length 2448;
	Best Local Similarity 88.5%; Pred. No. 0;	
	Matches 2227; Conservative 0; Mismatches 6; Indels 284; Gaps 2;	
QY	218	GACGAGGTTGGACGCTTCCCTTCCCCAGCAATAAATCTCTCGGAGATCACTGTAGAA 277
Db	83	GACCTGTTGGACCTTTCCTTCCAGCAATAAATCTCTCGGAGATCACTGTAGAA 142
QY	278	TTGTAGTGGATGAAAAATCAGGTGACGTGACACTGGGAAGATACACAGACACAGTGGAAACAG 337
Db	143	TTGTAGTGGATGAAAAATCAGGTGACGTGACACTGGGAAGATACACAGACACAGTGGAAACAG 202
QY	338	TGATTAAACAAGCTCAAGGTTCTTAAAGAAGACAGATGCCCTTTACAGACTGGGAGATGCA 397
Db	203	TGATTAAACAAGCTGAGGTTGTTAAGAAGACAGACATGCCCTTTACAGACTGGGAGATGCA 262
QY	398	TCTACTTGGTGTA CAGGAAGATGAA CCGGAACA CAACGTGGCATACCTCTATGAATCTT 457
Db	263	TCTACTTGGTGTA CAGGAAGATGAA CCGGAACA CAACGTGGCATACCTCTATGAATCTT 322
QY	458	TAAGTGAAGAACAAGGATGACACAAGAAATCCTTTGAAGCTAAACAAGGAAAATGTGTCC 517
Db	323	TAAGTGAAGAACAAGGATGACACAAGAAATCCTT----- 356
QY	518	ATGGGACCAAAAGATACCTCAGGTGCAAGGTGACGGCGAGGGCCGATCCCGGGTCCCTC 577
Db	357	----- 356
QY	578	CGTCGTGCGCCGCACTCAGGTGTGCTTTTGAGGAACCAACGCCATCAACATCGAGCTCAG 637
Db	357	----- 356
QY	638	ACCTCTTCCCACAGCCTCGCGCTCTTCCACGAGCCTTCTCTGACGGGGAGAGCGTT 697
Db	357	----- 356
QY	698	CCTCCAGTTGTGGGTCTGGGGTGTGGCATCTCCCTTAAAGGAAGTGGTCCCTCTGTGG 757
Db	357	-----TGGGTCTGGGGTGTGGCATCTCCCTTAAAGGAAGTGGTCCCTCTGTGG 406
QY	758	CAAGTGATGAAGTCTCCAGCTTTGGCTCAGCTCTCCAGACAGAAAGACTGGGTCCCTTT 817
Db	407	CAAGTGATGAAGTCTCCAGCTTTGGCTCAGCTCTCCAGACAGAAAGACTGGGTCCCTTT 466
QY	818	CGTCTGCGAAACCCAGGATCAGGAGGATTTGGAGCCCGTGAAGAAGAAAATGAGAGAG 877

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1. Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,
Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.

AUTHORS

Novel nucleic acids and polypeptides

TITLE

Patent: WO 0222660-A 57 21-MAR-2002;

JOURNAL

HYSEQ, INC. (US)

FEATURES

source

Location/Qualifiers

1..3181

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

96..1967

/notes="unnamed protein product"

/codon_start=1

/protein_id="CAD34749.1"

/db_xref="GI:21438724"

/db_xref="REVIEW:BL:CAD34749"

/translations="MERPEEGKQSPPPQWGRLLRLGABEGPHVLLRKREWTIGRRR
GCDLSPSNKLVGDHRIIVDKSQVLTLEDSTGTVINKLVKKYKOTCPULQTDV
IYLVYRNEPHNVAYLYLSKQKQWTFESFVPCVQAQGLKLLGSSDPTLASQ
SIVITGSGGGGIPKVGSPVADSVSSPASALPDRKTASFSLSPQDQEDLEPVKKK
MRGDDLDLNGQLLVQPRNAOTVHEDVRAAAGKDKMEETLCTICODLLHDCVSL
QPCWETFCACYSGWMSRLCTCPVERICKNHLNLYEAYLIQHPDKSRSED
YRQQAQPHCPGEPGPAQALGAPPTVSLITAVQDIYVCPLOGSHALCTCCFPQ
MPDRAREQDPVAPQCAVCIQPFCHLYGCTRTGCGLAPFCELMGLKDLGV
LNNYSYEDLKNLYATRGITWMLTESLVALQGVLLSYRVTVLVCYCGCLR
SPRELTVYQONIPASELPVAVTSRPDCYWGRCNRTQVKAHAKFNHI CEQTRFKN"

ORIGIN

Query Match 83.8%; Score 2244.6; DB 6; Length 3181;

Best Local Similarity 92.3%; Pred. No. 0;

Matches 2447; Conservative 0; Mismatches 79; Indels 125; Gaps 3;

```
QY 16 GCGCGCAATGTCCTTACACGCGCGCGCGCGAGCGGTTCCGGGTTCCGGGTTCCGGGCGGCGG 75
DB 21 GCGCGTACCGTCCGGAATCCCGGGTCCGACGATTTCTGTTCCGGGTTCCGGGCGGCGGCG 80
QY 76 GGGATGTAATCCGATGGAGCGGCCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 135
DB 81 GGGATGTAATCCGATGGAGCGGCCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 140
QY 136 TGGGCGCGCTCCTGCGTCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 195
DB 141 TGGGCGCGCTCCTGCGTCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 200
QY 196 CGGAGTGGACCATCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 255
DB 201 CGGAGTGGACCATCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 260
QY 256 GTCTCTGGATCACTGTAGAAATGTAGTGGATGAAATGAAATGAAATGAAATGAAATGAAATG 315
DB 261 GTCTCTGGATCACTGTAGAAATGTAGTGGATGAAATGAAATGAAATGAAATGAAATGAAATG 320
QY 316 GATACACGACCCAGTGGAACTGATTAACAAGCTGAAGTTGTTAAGAAGCAGACATGC 375
DB 321 GATACACGACCCAGTGGAACTGATTAACAAGCTGAAGTTGTTAAGAAGCAGACATGC 380
QY 376 CCTTTACAGACTGGGATGTCATCTACTGTTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 435
DB 381 CCTTTACAGACTGGGATGTCATCTACTGTTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 440
QY 436 GTGGCATACCTCTATGAATCTTTAAGTGAAGAAGCAGGATGACACAGAATCCTTTGAA 495
DB 441 GTGGCATACCTCTATGAATCTTTAAGTGAAGAAGCAGGATGACACAGAATCCTTTGAG 500
QY 496 GCTAACAGGAAATGTGTTCCATGGGACCAAGATACCTCAGTGCAGGTGCAGGCGGCGGA 555
```


Db 1858 TAACTGCCGACTCAGGTGAAGCTCACCACGCCATGAATTCATATCATATCTGTGAACA 1917
Qy 2067 GACAAGGTTCAAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTTTCAGCATGAGAGTGAA 2126
Db 1918 GACAAGGTTCAAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTTTCAGCATGAGAGTGAA 1977
Qy 2127 GAGAGCGTGTGTTTAAATAACAGAGCAAGCAGTCAAGGTGTTTTCACAGCCCCCTGAG 2186
Db 1978 GAGAGCGTGTGTTTAAATAACAGAGGAGCAGCAGTCAAGGTGTTTTCACAGCCCCCTGAG 2037
Qy 2187 GGAAGGAGCGCAGGCTTCCGACAGAGTGTCTGAGGAGTCACTCTCTGAGAGCTTTTAC 2246
Db 2038 GGAAGGAGCGCAGGCTTCCGACAGAGTGTCTGAGGAGTCACTCTCTGAGAGC--TTTAC 2095
Qy 2247 CCTCTGAGTGAGACCTCCCGAGAGCCCGGAGCGCCGAGCGCCCTCTCTGAGAGCGC 2306
Db 2096 CCTCTGAGTGAGACCTCCCGAGAGCCCGGAGCGCCGAGCGCCCTCTCTGAGAGCGC 2155
Qy 2307 TGGCAGGCGCTCGTGTGGCATACAGCAGCAGAGCGAAGCCTTCTGTAACATCGCGCGC 2366
Db 2156 TGGCAGGCGCTCGTGTGGCATACAGCAGCAGAGCGAAGCCTTCTGTAACATCGCGCGC 2215
Qy 2367 TCCCGCCGAGAGGCGCAGTGTGCTCTTCTGATCTTCCGAACTACAGTTAAAGCAG 2426
Db 2216 TCCTGCCGAGAGGCGCAGTGTGCTCTTCTGATCTTCCGAACTACAGTTAAAGCGG 2275
Qy 2427 AAGTCTGTGTTTCAAGAAAGTTTCAAGGAGAGCGCAAGTTTATCAAAACATGTTTC 2486
Db 2276 AAGTCTGTGTTTCAAGAAAGTTTCAAGGAGAGCGCAAGTTTATCAAAACATGTTTC 2335
Qy 2487 AGGAGAGGAGCATAAGTTTACAGCTACAGCGTACACAATATCTGCTGCTGGAA 2546
Db 2336 AGGAGAGGAGCATAAGTTTACAGCTACAGCGTACACAATATCTGCTGCTGGAA 2395
Qy 2547 AACACAGCATTATCTATCTATTTTATTTAATAGTGTGCTGCTATCTTCTAATAAGA 2606
Db 2396 AACACAGCATTATCTATCTATTTTATTTAATAGTGTGCTGCTATCTTCTAATAAGA 2455
Qy 2607 TTTAAATGTCACAACTGTACACAAATAATATATTTATTAATTTTACAAATGACAAA 2666
Db 2456 TTTAAATGTCACAACTGTACACAAATAATATATTTATTAATTTTACAAATTTGACTAAA 2515

RESULT 8
AK001658 3138 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ10796 fis, clone NT2RP400648, weakly similar
DEFINITION to TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0.
ACCESSION AK001658
VERSION AK001658.1 GI:7023050
KEYWORDS oligo capsing; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y.,
Matsunawa, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
Wakamatsu, A., Nakamura, Y., Negahari, K., Masuho, Y. and Sasaki, N.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3138)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1332-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center

etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Source
Location/Qualifiers
1..3138
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP400648"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP4"
/note="cloning vector: pME18FL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
CDS
65..1936
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91817.1"
/db_xref="GI:7023051"
/translation="MERPEEGKQSPPPWGLLLRLGAEGBEPHVLRLKREWTIGRRR
GDLFPFNKLVSGDHCRIVVDEKSGQVLTSTSGTVINKLVKKQKTCPLQTDV
IYLVTRKNPEHNVAIYLSSEKQMTQESFEMVPCVQAAGLKLSSDPTLAAQ
ILVITRSGGGISPKGSPSVASDEVSPASALPDRKTASFSLSPQQLDEPKKK
MRGDDLDLNGOLLVAQPRNACTVHEDVRAAGPKDMEETLCIICQDLHDCVSL
QCMHTFCACSGHMERSSLCPTCPCVERICKHILNNLVEALIHPDKSRSEED
VSMARNKITQMLQPKVRVSFSEEGSEDLLELSDVDSSESDISQYVVCRCPE
YRRAAQPHCPAPEGEPAGALGDAPPTSVLSLTTAVQDYVCPQGGHALCTCCFP
MDPRAERQDPRVAPQCAVCLQPFCHLYNGCTRTGTCYCLAPFCEINLGDKLDGV
LNNNSVESDILKNYLATRLTKWMLTSLVALORGVFLISDYRTVGLCYCCGLR
SPRELYQYQONIPASELPVAVTSRDPDCYWGRCNRTQVKAHMKFNHICEQTRFN"

ORIGIN
Query Match 84.4%; Score 2262.4; DB 9; Length 3138;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 2454; Conservative 0; Mismatches 61; Indels 125; Gaps 3;
Qy 27 CTCTTCACAGCGCGCGCGCGAGCGGTTCCGGGTTCCGGCGCGCGCGCGCGGATGGAAT 86
Db 1 CTCTTCACAGCGCGCGCGCGAGCGGTTCCGGGTTCCGGGTTCCGGCGCGCGCGGATGGAAT 60
Qy 87 CCCGATGAGCGCGCGCGAGGAGGCAAGCAGTCGCGCGCGCGCGCGCGCGCGCGCT 146
Db 61 CCCGATGAGCGCGCGCGAGGAGGCAAGCAGTCGCGCGCGCGCGCGCGCGCGCGCT 120
Qy 147 CTTGCTCTGCGCGCGCGAGGAGGCGGAGCGCGAGCTCTCTGAGGAGCGGAGTGGAC 206
Db 121 CTTGCTCTGCGCGCGCGAGGAGGCGGAGCGCGAGCTCTCTGAGGAGCGGAGTGGAC 180
Qy 207 CATCGGCGGAGACGAGGTTGCGACCTTTCTTCCCGAGCAATAAACTGCTCTTGAGA 266
Db 181 CATCGGCGGAGACGAGGTTGCGACCTTTCTTCCCGAGCAATAAACTGCTCTTGAGA 240
Qy 267 TCACTGTAGATTGTAGTCGATGAAAATCAGTCAAGTTCAGTTCGAGTTCGAGTACCGAC 326
Db 241 TCACTGTAGATTGTAGTCGATGAAAATCAGTCAAGTTCAGTTCGAGTTCGAGTACCGAC 300
Qy 327 CAGTGAACACAGTGATTAAACAAGCTGAAGGTTGTTAAGAAAGCAGACATGCCCTTTACAGAC 386
Db 301 CAGTGAACACAGTGATTAAACAAGCTGAAGGTTGTTAAGAAAGCAGACATGCCCTTTACAGAC 360
Qy 387 TGGGATGTCATCTACTTGTGTACAGGAAGATGACCCGGAACACAACTGCGCATCT 446
Db 361 TGGGATGTCATCTACTTGTGTACAGGAAGATGACCCGGAACACAACTGCGCATCT 420
Qy 447 CTATGAATCTTTAAGTGAAGCAAGGCTGATGACAAAGAACTCTTTGAAGCTAAACAAGGA 506
Db 421 CTATGAATCTTTAAGTGAAGCAAGGCTGATGACAAAGAACTCTTTGAAGCTAAACAAGGA 480
Qy 507 AATATGTTCCATGGGACCAAGATATCTCAGGTGAGGTGAGGCGGAGCGGCGCATCC 566
Db 481 CTGTGTTGCGCAGGCTGCTTAAACTCTCTGGGATCAAGT----- 520
Qy 567 CCGGGTCCCTCGTCTCGCCCGCACTCAGGTGTGCTTTGAGGAACACACAGCATCAAC 626

[illegible]

778	AAATGCCAAA	CCGTCCACGAGGAGCTCAGAGCAGCGGCTGGGAGCCAGACAGATGCA	837
987	GGAGACGCTGAC	CATGATCATCTGCCAGGACCTGCTGCAAGCATGCGTAGGTTTGACGC	1046
838	GGAGACGCTGAC	CATGATCATCTGCCAGGACCTGCTGCAAGCATGCGTAGGTTTGACGC	897
1047	CTGCAATGCA	CACGTTCTGGCGGCTTTGCTACTCGGGCTGGATGAGAGCGCTCGTCCCTGTG	1106
898	CTGCATGCA	CACGTTCTGGCGGCTTTGCTACTCGGGCTGGATGAGAGCGCTCGTCCCTGTG	957
1107	TCCTA	CCTGCGCTGTCCGCTGGAGGGATCTGTAAAAACCAATCCTCAACACCTCGT	1166
958	TCCTA	CCTGCGCTGTCCGCTGGAGGGATCTGTAAAAACCAATCCTCAACACCTCGT	1017
1167	GGAGCAT	ACCTCATCCAGCATCCAGACAGAGCTCGCAGTGAAGAAGATGTGCAAAAGTAT	1226
1018	GGAGCAT	ACCTCATCCAGCATCCAGACAGAGCTCGCAGTGAAGAAGATGTGCAAAAGTAT	1077
1227	GGATGCCAGGAA	TAAATCACTCAAGACATGCTGCAGCCCAAAAGTCAGGCGGCTTTTTTC	1286
1078	GGATGCCAGGAA	TAAATCACTCAAGACATGCTGCAGCCCAAAAGTCAGGCGGCTTTTTTC	1137
1287	TGATGAAGAAGGAG	TTTCAAGAGACCTGCTGGAGCTGCAGAGTTGACAGTGAAGTGAAGTCTTC	1346
1138	TGATGAAGAAGGAG	TTTCAAGAGACCTGCTGGAGCTGCAGAGTTGACAGTGAAGTGAAGTCTTC	1197
1347	AGACATTTAGCCAGC	CAATACGTCGTGTGTCGCCGAGTGTCTTGATACAGAAGCAGAGCGGC	1406
1198	AGACATTTAGCCAGC	CAATACGTCGTGTGTCGCCGAGTGTCTTGATACAGAAGCAGAGCGGC	1257
1407	GCAGCTTCCCA	CTGCCAGCACTCGAGAGGCGAGCCAGAGCCCAAGGCCCTGGGGGA	1466
1258	GCAGCTTCCCA	CTGCCAGCACTCGAGAGGCGAGCCAGAGCCCAAGGCCCTGGGGGA	1317
1467	TGCACCTTCCA	CGTTCGCTGAGCTGACGACAGCAGTCCAGAGTTACGTGTGCCCTCTGCA	1526
1318	TGCACCCCCAC	GTTCGCTGAGCTGACGACAGCAGTCCAGAGTTACGTGTGCCCTCTGCA	1377
1527	AGGAAGCCACG	CCCTGTGCACTGCTGCTTCCAGCCCATCCCGACGGAGAGCGAGCG	1586
1378	AGGAAGCCACG	CCCTGTGCACTGCTGCTTCCAGCCCATCCCGACGGAGAGCGAGCG	1437
1587	CGAGCAGGAC	CCCGGTGTGCGCCCTCAGCAGTGTGGGTGTGCTGCTGACGAGCTTTTGCCA	1646
1438	CGAGCAGGAC	CCCGGTGTGCGCCCTCAGCAGTGTGGGTGTGCTGCTGACGAGCTTTTGCCA	1497
1647	CCTGTACTGGG	CTGACCCGAGCCGGCTGTACGGCTGCTGGCCCGCTTTTGTGAGCT	1706
1498	CCTGTACTGGG	CTGACCCGAGCCGGCTGTACGGCTGCTGGCCCGCTTTTGTGAGCT	1557
1707	CAACCTGGGTG	ACAAAGTGTCTGGAACGGCTGTCTGAACAAACAAGCTTACAGTACAGCAT	1766
1558	CAACCTGGGTG	ACAAAGTGTCTGGAACGGCTGTCTGAACAAACAAGCTTACAGTACAGCAT	1617
1767	CCTGAGAAATTA	CCTGCGCAACAGAGTTTGACATGGAAAAACAATGTTGACCGAGAGCCT	1826
1618	CCTGAGAAATTA	CCTGCGCAACAGAGTTTGACATGGAAAAACAATGTTGACCGAGAGCCT	1677
1827	CGTGGCTCTC	CAGCGGGAGTGTTCCTGCTCTTGATTAACAGAGTACGGAGACACCGT	1886
1678	CGTGGCTCTC	CAGCGGGAGTGTTCCTGCTCTTGATTAACAGAGTACGGAGACACCGT	1737
1887	TCGTGTACTG	CTGTGGCCCTGCGACGCTTCCGCTGAGCTGACCTTACAGTATCGGCAGAA	1946
1738	TCGTGTACTG	CTGTGGCCCTGCGACGCTTCCGCTGAGCTGACCTTACAGTATCGGCAGAA	1797
1947	CATTTCCTGCT	TCCAGTGTGCGAGTGGCCGTAAACAATCCCGCTGCTGACTGTACTGGGGCGG	2006
1798	CATTTCCTGCT	TCCAGTGTGCGAGTGGCCGTAAACAATCCCGCTGCTGACTGTACTGGGGCGG	1857
2007	TAACTGCGCA	CTCAGGTGAAGCTCACACGCCATGAATTCATCATATCTGTGAACA	2066
1858	TAACTGCGCA	CTCAGGTGAAGCTCACACGCCATGAATTCATCATATCTGTGAACA	1917

RESULT 7

BD156548

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

REFERENCES

AUTHORS

TITLE

BD156548 3138 bp DNA linear PAT 17-JAN-2003

Primer for synthesizing full-length cDNA and use thereof.

BD156548

BD156548.1 GI:27862306

JP 2002191363-A/11391.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Catarrhini: Hominoidea: Hominidae: Homininae: Homo.

Mammalia; Euthera; Filices; Cataglyph; Homoptera; Homoc.

I (Dares 1 CO 3138)
Ota, T., Tsogai, T.: Nishikawa, T.: Hayashi, K.: Saito, K.: Yamamoto, J.:

Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002191363-A 11391 09-JUL-2002;

HELIUM RESEARCH INSTITUTE

OS Homo sapiens (human)

PN · JP 2002191363-A/11391

PD 09-JUL-2002

28-JUL-2000 JP 2000280990

PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU

PI SAI TO,
TAKASHI YAMAMOTO
AT WAKAMATSU

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AT WAKAMATSU,
DI KEIICHI NAGAI TETSUJI OTSUKI

PT
PC
NETTCHI NAGAT, IETSUOTI OISUKI

PC
C12N15/09 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/PC

10.

PC C12P21/02.C12O1/68//C12P21/08.G06F17/30.C12N15/00.C12N5/00 CC

misc feature

```

/clone_11b="NIH_MGC_21"
/lab_host="DH10B-8"
/notes=Vector: POTB7"
1. 3189
   /gene="CHFR"
/notes= synonyms: FLJ10796, RNF116"
/db xref="LocusID:55743"
/db xref="MIM:605209"
9. 1967
/codon_start=1
/products="CHFR protein"
/protein_id="AAH12072.1"
/db xref="GI:15082330"
/db xref="LocusID:55743"
/translation="MERPEQKQSPPOPGWRLNLGABEGEPHVLLRKWEWIGRRR
GCDLSPFNKLVGSDHRIVDCEQVTLSDTSGTVINKLKVKKQKCPGLGTGVV
IYLVKRNTEFNHVAIYLSSEKQMTQEFDSFGAGAGAPRPVPPSPATQVCF
EEQPTSTFSDLPFTASASPEFSPAGBERSSCSGGGGSPKSGSPVASDEVFS
ASALPDKRTASFSLEQQDLLEPVKKMGDDGLDNGQLLVAPRNATQVHEDV
RAAAGQKDMETLTCIIQQDLHDVCVLCQMFTCAACYSGHWMERSSLCPTCRCPV
RETCKNHLNIAVEAYLIQHPDKSSSEEDVSDMARNKITODMLQPKVRASFGEDEGS
SEDLLESDVDSESDIYQPVVVCQCEYRQAAQPHCPABEGFGAQLGDAP
TSVSLTATGVQVCLPQGSHALCTCCFQPMRVERODPRPAQOCVCLPFCFL
YMGCTTGYGCVCLAPCELNALGDCFLQVLNNVSIEDILKNYLTSLGTLKMLMLES
LVALQSGVFLSDYRVCTGDTVLCYCCGLARSFRLTYQYRQNI PASELPVAVTSPDCY
WGRNCTGVQLKXHHAKFNHCQGTFRKN"
120. 317
/notes=FFA; Region: FFA domain. The FFA
(Forkhead-associated) domain is a phosphopeptide binding
motif"
/db xref="CDD:pfam00498"
879. 1010
/notes=HRD1; Region: COG5243, HRD1, HRD ubiquitin ligase

```

	Query Match	95.8%;	Score 2566.4;	DB 9;	Length 2639;
	Best Local Similarity	98.6%;	Pred. No. 0;		
	Matches 2613;	Conservative	0; Mismatches	1; Indels	36; Gaps
QY	17	GCGCAATGTCTTTGACAGCGCGGGCGGCAGCCGGTTCGGGTTCGGCGCGGGCGGG	76		
DB	5	GCCGCAATGTCTTTGACAGCGCGGGCGGCAGCCGGTTCGGGTTCGGCGCGGGCGGG	64		
QY	77	GGATGTGAATCCCATGTAGAGCGGCCCGAGGAGGCAAGCAGTTCGCGCGCGCGGACGCT	136		
DB	65	GGATGTGNATCCCATGTAGAGCGGCCCGAGGAGGAGGCAAGCAGTTCGCGCGCGCGGACGCT	124		
QY	137	GGGACCGGCTCCTTGCTCTGGCGCGGAGGAGGCGGAGCGGACCGGACGTCCTCTGTAGGAAGC	196		
DB	125	GGGACCGGCTCCTTGCTCTGGCGCGGAGGAGGCGGAGCGGACCGGACGTCCTCTGTAGGAAGC	184		
QY	197	GGGAGTGGACCATTGGGCGGAGACGAGGTTCGACCTTTCTTCCCACGAATAAATCTGG	256		
DB	185	GGGAGTGGACCATTGGGCGGAGACGAGGTTCGACCTTTCTTCCCACGAATAAATCTGG	244		
QY	257	TCTCTGGAGATCACGTGTAGAATTTGATGGATGAAAAATACAGGTCAAGTGTACATGGAAAG	316		
DB	245	TCTCTGGAGATCACGTGTAGAATTTGATGGATGAAAAATACAGGTCAAGTGTACATGGAAAG	304		
QY	317	ATACGACACAGTGGACATGATTAACAAGCTGAAGTTGTTAAAGAACAGACATGCTCC	376		
DB	305	ATACGACACAGTGGACATGATTAACAAGCTGAAGTTGTTAAAGAACAGACATGCTCC	364		
QY	377	CTTTACAGACTGGGGATGTCTACTCTTGGTGTACAGAAAGAAATGAACCGGAAACACAACG	436		
DB	365	CTTTACAGACTGGGGATGTCTACTCTTGGTGTACAGAAAGAAATGAACCGGAAACACAACG	424		
QY	437	TGGCATACCTCTATGAATCTTTAAGTGAAGAACAGGATCACACAGAAATCTTTGAG	496		
DB	425	TGGCATACCTCTATGAATCTTTAAGTGAAGAACAGGATCACACAGAAATCTTTGAG	480		
QY	497	CTAACAGGAAAATGTGTTCCATGGGACCAAAGATACCTCAGGTGTCAGGTTCAGGGGAG	556		
DB	481	-----GATACCTCAGGTGTCAGGTTCAGGGGAG	508		
QY	557	GGGCGGATPCCCGGGTCCCTCGTGTGCGCCGCACTCAGGTGTGCTTTGAGGAAACAC	616		
DB	509	GGGCGGATPCCCGGGTCCCTCGTGTGCGCCGCACTCAGGTGTGCTTTGAGGAAACAC	568		
QY	617	AGCCATCAACATCGACGTACAGCTCTTCCCAACAGCCTCGGCTCTTCCACCGAGGCTT	676		
DB	569	AGCCATCAACATCGACGTACAGCTCTTCCCAACAGCCTCGGCTCTTCCACCGAGGCTT	628		
QY	677	CTCTGCAAGCGCGAGAGCGTTCTCCAAGTGTGGGTCTGCGGGTGGTGGCATCTCCCTTA	736		
DB	629	CTCTGCAAGCGCGAGAGCGTTCTCCAAGTGTGGGTCTGCGGGTGGTGGCATCTCCCTTA	688		
QY	737	AAGGAAGTGTCCCTCTGTGCGAAGTGAATGAAGTCTCCAGCTTTGCGCTTCAGCTCTCCAG	796		
DB	689	AAGGAAGTGTCCCTCTGTGCGAAGTGAATGAAGTCTCCAGCTTTGCGCTTCAGCTCTCCAG	748		
QY	797	ACAGAAAGACTGCGTCTTTTTGCTGTTGGAACCCACAGATCAGGAGGATTTGGAGCCCG	856		
DB	749	ACAGAAAGACTGCGTCTTTTTGCTGTTGGAACCCACAGATCAGGAGGATTTGGAGCCCG	808		
QY	857	TGAAGAGAAATGAGAGGATGCGGACCTTGACCTGAACCGGCGAGTTGTTGCTCGAC	916		
DB	809	TGAAGAGAAATGAGAGGATGCGGACCTTGACCTGAACCGGCGAGTTGTTGCTCGAC	868		
QY	917	AACCGCGTAGAAATGCCCAAAACCGTCCAACGAGGACGTCAGAGACGCGCTGGGAAGCCAG	976		
DB	869	AACCGCGTAGAAATGCCCAAAACCGTCCAACGAGGACGTCAGAGACGCGCTGGGAAGCCAG	928		
QY	977	ACAAGATGGAGAGACGCTGACATGTCATCTGCCAGGACCTGCTGCACGACTGCGTGA	1036		
DB	929	ACAAGATGGAGAGACGCTGACATGTCATCTGCCAGGACCTGCTGCACGACTGCGTGA	988		

QY	1037	GTTTGCAGCCCTGCATGCACACGTTCTGCGCGCTTCTACTCTCGGCTCGATCGAGCGCT	1099
DB	989	GTTTGCAGCCCTGCATGCACACGTTCTGCGCGCTTCTACTCTCGGCTCGATCGAGCGCT	1048
QY	1097	CGTCCCTGTCTCTACTCTGCGGTGTCCTGCTGAGCGGATCTCTAARAAACCATCCTCA	1156
DB	1049	CGTCCCTGTCTCTACTCTGCGGTGTCCTGCTGAGCGGATCTCTAARAAACCATCCTCA	1108
QY	1157	ACAACCTCGTGGAGACATACCTCATCCAGCATCCAGACAAGAGTTCGCAAGAAAGATG	1216
DB	1109	ACAACCTCGTGGAGACATACCTCATCCAGCATCCAGACAAGAGTTCGCAAGAAAGATG	1168
QY	1217	TGCAAGTATGATGCCAGGAATAAATCACTCAAGCATGCTGCAGCCCAAGTCAGGC	1276
DB	1169	TGCAAGTATGATGCCAGGAATAAATCACTCAAGCATGCTGCAGCCCAAGTCAGGC	1228
QY	1277	GCTCTTTTTCTGATGAAGAGGAGTTTCAGAGCACCTGCTGGAGCTGTTCAGAGCTTGACA	1336
DB	1229	GCTCTTTTTCTGATGAAGAGGAGTTTCAGAGCACCTGCTGGAGCTGTTCAGAGCTTGACA	1288
QY	1337	GTGAGTCTCTCAGACATTAGCCAGGCATACGTCTGTGTGCCGCGAGTCTCTGAGTACAGAA	1396
DB	1289	GTGAGTCTCTCAGACATTAGCCAGGCATACGTCTGTGTGCCGCGAGTCTCTGAGTACAGAA	1348
QY	1397	GGCAGGCGGGCAGCCTCCCACTGCCACGACCCGAGGCGAGCCAGGAGCCCCACAGG	1456
DB	1349	GGCAGGCGGGCAGCCTCCCACTGCCACGACCCGAGGCGAGCCAGGAGCCCCACAGG	1408
QY	1457	CCCTGGGGATGCACCTTCCACGTCGTCTGAGCTGACGACAGCAGTCCAGGATTAAGTGT	1516
DB	1409	CCCTGGGGATGCACCTTCCACGTCGTCTGAGCTGACGACAGCAGTCCAGGATTAAGTGT	1468
QY	1517	GCCCTCTGCAGGAAGACGACCCCTGTGCACTGCTGCTTCCAGCCCATGCCCGACCGGA	1576
DB	1469	GCCCTCTGCAGGAAGACGACCCCTGTGCACTGCTGCTTCCAGCCCATGCCCGACCGGA	1528
QY	1577	GAGCGGAGCGGACGAGGACCCCGTGTGCCCTCTCAGCAGTGTGCGGTGCTGCTGCAGC	1636
DB	1529	GAGCGGAGCGGACGAGGACCCCGTGTGCCCTCTCAGCAGTGTGCGGTGCTGCTGCAGC	1588
QY	1637	CTTTCTGCCACTGTACTTGCGGTGCACCCGACCCGCTGCTACGGCTGCCTGCCGCCGT	1696
DB	1589	CTTTCTGCCACTGTACTTGCGGTGCACCCGACCCGCTGCTACGGCTGCCTGCCGCCGT	1648
QY	1697	TTTGTGAGCTCAACCTGGGTGACAAGTGTCTGACGCGCGTGTCTGAAACAAACAGACTACG	1756
DB	1649	TTTGTGAGCTCAACCTGGGTGACAAGTGTCTGACGCGCGTGTCTGAAACAAACAGACTACG	1708
QY	1757	AGTCAGACATCTCGAAGAATTACTGTGGAAACAGAGGTTTGACATGGAAAAACATGTTGA	1816
DB	1709	AGTCAGACATCTCGAAGAATTACTGTGGAAACAGAGGTTTGACATGGAAAAACATGTTGA	1768
QY	1817	CCGAGACCTCTGTGGCTCTCCAGCGGGAGTGTTCTGCTGTCTGATTCAGAGTTCACGG	1876
DB	1769	CCGAGACCTCTGTGGCTCTCCAGCGGGAGTGTTCTGCTGTCTGATTCAGAGTTCACGG	1828
QY	1877	GAGACACCGTTCTGTGTTACTGTGTGGCCTGCGCAGCTTTCGTGTAGCTGACCTATCAGT	1936
DB	1829	GAGACACCGTTCTGTGTTACTGTGTGGCCTGCGCAGCTTTCGTGTAGCTGACCTATCAGT	1888
QY	1937	ATCGGCAGAACATTCCTGCTTCCGAGTTGCCAGTGCAGTCGCTAACATCCGTCCTGACTGCT	1996
DB	1889	ATCGGCAGAACATTCCTGCTTCCGAGTTGCCAGTGCAGTCGCTAACATCCGTCCTGACTGCT	1948
QY	1997	ACTGGGCCCGTAACTGCGCGCACTCAGGTGAAAGCTCACCACGCCATGAAATTCAAATCAT	2056
DB	1949	ACTGGGCCCGTAACTGCGCGCACTCAGGTGAAAGCTCACCACGCCATGAAATTCAAATCAT	2008
QY	2057	TCTGTGAACAGACAGGTTCCAAAACTTAAGCATCCAGAGGCCCTGTAGCAGCTTTCAGCAC	2116
DB	2009	TCTGTGAACAGACAGGTTCCAAAACTTAAGCATCCAGAGGCCCTGTAGCAGCTTTCAGCAC	2068
QY	2117	TGGAGGTGAAGAGAGCGGTGTTTTTAAATAACAGAGACAAGCAGCTCAAGGTGTTTTTACA	2176

Db	1409	CCCTGGGGATGCA	CCCTTCCAGTCCGTCAGCTGACGACAGCAGTCCAGATTACGTGT	1468
Qy	1517	GCCTCTGCAAGAAAGCCAC	CGCCCTGTGTGACCTCTTCAGAGCCCAATGCCCGACCGGA	1576
Db	1469	GCCCTCTGCAAGAAAGCCAC	CGCCCTGTGTGACCTCTTCAGAGCCCAATGCCCGACCGGA	1528
Qy	1577	GACGGAGCCGACAGCAGAC	CCGCGTGTCCGCTCAGCAGTGTGGGTCTGCTCTGCAGC	1636
Db	1529	GACGGAGCCGACAGCAGAC	CCGCGTGTCCGCTCAGCAGTGTGGGTCTGCTCTGCAGC	1588
Qy	1637	CTTTCTGCCACCTGTACT	CTGGGGTGTACACCGGACCGGCTGCTAGGCTGCTGGCCCCGT	1696
Db	1589	CTTTCTGCCACCTGTACT	CTGGGGTGTGTACACCGGACCGGCTGCTAGGCTGCTGGCCCCGT	1648
Qy	1697	TTTTGTAGCTCAACCTGGGT	GACAAAGTGTGAGACGGCGTGTGAACAAACAGACTACG	1756
Db	1649	TTTTGTAGCTCAACCTGGGT	GACAAAGTGTGTGAGACGGCGTGTGAGAACAAACAGACTACG	1708
Qy	1757	AGTCAGACATCTCTGAAGAA	TACCTGTGCAACCGAGAGTTTGACATGAAAAACAATGTGA	1816
Db	1709	AGTCAGACATCTCTGAAGAA	TACCTGTGCAACCGAGAGTTTGACATGAAAAACAATGTGA	1768
Qy	1817	CCGAGACCTCTGTGGCTCT	CCAGCGGGAGTGTTCGTCTGTCTGATTACAGAGTACAGG	1876
Db	1769	CCGAGACCTCTGTGGCTCT	CCAGCGGGAGTGTTCGTCTGTCTGATTACAGAGTACAGG	1828
Qy	1877	GAGACACCGTTCTGTGTACT	GTGTGGCTGTGCGCAGCTTCCGCTGAGCTGACCTATCAGT	1936
Db	1829	GAGACACCGTTCTGTGTACT	GTGTGGCTGTGCGCAGCTTCCGCTGAGCTGACCTATCAGT	1888
Qy	1937	ATGGGAGAAATTCCTGCTT	CCGAGTTCGCAAGTGGCGGTAAACATCCGCTCTGCAGTCT	1996
Db	1889	ATGGGAGAAATTCCTGCTT	CCGAGTTCGCAAGTGGCGGTAAACATCCGCTCTGCAGTCT	1948
Qy	1997	ACTGGGCGCGTAACCTGCG	CACTCAGGTGAAAGTCCACCGCATGAAATTCAAATCAT	2056
Db	1949	ACTGGGCGCGTAACCTGCG	CACTCAGGTGAAAGTCCACCGCATGAAATTCAAATCAT	2008
Qy	2057	TCTGTGAACAGAAAGTTT	CAAAACTTAAGCATCCAGAGGCGCTGAGCAGTTTCAGCAC	2116
Db	2009	TCTGTGAACAGAAAGTTT	CAAAACTTAAGCATCCAGAGGCGCTGAGCAGTTTCAGCAC	2068
Qy	2117	TGGAGGTGAAGAGACGTGT	TTTTTAAATAACAGAGACAGCACGCTCAAGGTGTTTTTCA	2176
Db	2069	TGGAGGTGAAGAGACGTGT	TTTTTAAATAACAGAGACAGCACGCTCAAGGTGTTTTTCA	2128
Qy	2177	GCCCCCTGAGGGAAGGAG	CGAGGCTCTCCAGAGGTCTCTGGGTGACCTCTCTGTGG	2236
Db	2129	GCCCCCTGAGGGAAGGAG	CGAGGCTCTCCAGAGGTCTCTGGGTGACCTCTCTGTGG	2188
Qy	2237	AGCTTTTACCTCTGAGT	GAGACCTCCCGAGAGCCCGGGGCGCGAGCCCGCCCTCC	2296
Db	2189	AGCTTTTACCTCTGAGT	GAGACCTCCCGAGAGCCCGGGGCGCGAGCCCGCCCTCC	2248
Qy	2297	TGGTGAGCGCTGGCGAG	GGCTCGTGGTGCATCAGCAGCAGACGAGTCTTCTGTAA	2356
Db	2249	TGGTGAGCGCTGGCGAG	GGCTCGTGGTGCATCAGCAGCAGACGAGTCTTCTGTAA	2308
Qy	2357	CATCGCGCGTCCCGCGAG	GGGCGAGTTTGTCTCTTTTCTATATTTCCGAACTACA	2416
Db	2309	CATCGCGCGTCCCGCGAG	GGGCGAGTTTGTCTCTTTTCTATATTTCCGAACTACA	2368
Qy	2417	GTTTAAAGCAGAAGTCTGT	TTTTTTCAGGAAAGTTTCAAGGGAGAGGCAAGTTTATCAAAA	2476
Db	2369	GTTTAAAGCAGAAGTCTGT	TTTTTTCAGGAAAGTTTCAAGGGAGAGGCAAGTTTATCAAAA	2428
Qy	2477	ACATTTTTCAGGAGAGG	GAGCAATAGTTTACAGCTACAGGAGCTACACAATATCCTG	2536
Db	2429	ACATTTTTCAGGAGAGG	GAGCAATAGTTTACAGCTACAGGAGCTACACAATATCCTG	2488
Qy	2537	CTGCTGGGAAAAACAC	ACATTTTCTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	2596

```

Db      2489  CTGCTGGGAAAAACACAGCAGTATTTTCTATCTTTTATTTTAAATAGGTTTGGTGTCTAICT 2544
Qy      2597  TCTAATAAGATTAAATGTACAACTGTAGCACAATAATATAATTTATATAATTTACAAA 2656
Db      2549  TCTAATAAGATTAAATGTACAACTGTAGCACAATAATATAATTTATATAATTTACAAA 2608
Qy      2657  TTGACAAAAA 2666
Db      2609  TTGACTAAAA 2618

RESULT 4
AK027687
LOCUS   Homo sapiens cDNA FLJ14781 fis, clone NT2RP4000455, weakly similar
        to TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0.
ACCESSION
VERSION  AK027687.1 GI:14042552
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
        Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y.,
        Matsunawa,H., Ishii,S., Kawai,Y., Saito,K., Yamanoto,J.,
        Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuho,Y. and Sasaki,N.
        NEDO human cDNA sequencing project
TITLE    Unpublished
JOURNAL  2 (bases 1 to 2639)
REFERENCE Isogai,T. and Otsuki,T.
          Direct Submission
TITLE    Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
          Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0912, Japan
          (E-mail:genomics@hri.co.jp, Tel.81-438-52-3975, Fax:81-438-52-3986)
COMMENT  NEDO human cDNA sequencing project supported by Ministry of
          Economy, Trade and Industry of Japan; cDNA full insert sequencing:
          Research Association for Biotechnology; cDNA library construction,
          5'- & 3'-end pass sequencing and clone selection; Helix
          Research Institute (supported by Japan Key Technology Center etc.)
          and Department of Virology, Institute of Medical Science,
          University of Tokyo.
FEATURES
          Location/Qualifiers
          1..2639
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="NT2RP4000455"
             /cell_line="NT2"
             /cell_type="teratocarcinoma"
             /clone_lib="NT2RP4"
             /note="Cloning vector: pME18SFL3-mRNA from NT2 neuronal
             precursor cells after 2-weeks retinoic acid (RA)
             induction."
          79..2037
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="BAB55297.1"
             /db_xref="GI:14042553"
             /translation="MERPEEKQSPPPQPPQGRLLRLGAEGEHPHLRLKREWTIGRRRE
             GCDLSPFNKLNSGDHCRIVDEKSGVTLDETSTGTVINKLVKVKQTCPLOTQGV
             ILYLVKRNKEHNVAIYVLSLEKQGTQSFDTSGAGRGADPRVPPSPGATQVC
             EEPOPTSTDLFFTAGASTTEPSPAGRERSSCSGSGGGISPKSGSPSVASDVSS
             ASALPDKTASFSSLEPQDOEDLEPVKKVKGKDGDLINGOLLVAQPRNAQTVHED
             RAAGKDPKMEETLCTLICDLDHCVSLQPCMHTFCAACYCYSGWVRSLLCPTCRCPK
             ERICKNHLNLVEAYLILQHPDKRSERSEVDQMDARNKITQDMLQPKVRSFSDEGS
             SEDLELSDVDSSESDLSQPIVVCRCQCFEYRRAQAQPPHCPAGEGAPALGADAE
             TSVGLTAVDQVCPGLQSHALCTCCFPMDRAEREDQDPRAVQCAVCLQPFCHRE
             YWGTRTQCYGVCGLQFCELNGLDKDGLVNLRNSESIEDLKNYLATGLTKWMLTCE
             LVALQGVGTGDTLVLCYCCGLRSFRELTYQYRQNPISALPELFAVYTFEDPCY
             WGRNCRGTVCVAAHAKMKNFHCIOETFRFN"

```

ORIGIN

```

Qy 2657 TTGCAAAAA 2666
Db 2609 TTGACTAAAA 2618

RESULT 3
BD156534
LOCUS 2639 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD156534
VERSION BD156534.1 GI:27862292
KEYWORDS JP 2002191363-A/11377.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2639)
Ota,T., Isozaki,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patient: JP 2002191363-A 11377 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/11377
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOZAKI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
FT CDS Location/Qualifiers
1..2639
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 95.8%; Score 2566.4; DB 6; Length 2639;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2613; Conservative 0; Mismatches 1; Indels 36; Gaps 1;

Qy 17 GCCGCAATGCTCTTTGACAGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 76
Db 5 GCCGCAATGCTCTTTGACAGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 64

Qy 77 GGATGTGAATCCGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 136
Db 65 GGATGTGAATCCGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 124

Qy 137 GGGGACGCTCTCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 196
Db 125 GGGGACGCTCTCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 184

Qy 197 GGGAGTGAACCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 256
Db 185 GGGAGTGAACCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 244

Qy 257 TCTCTGGAGTACCTGTGAGATTGTAGTGGATTGAAATCAGGTGACATCTGGAAG 316
Db 245 TCTCTGGAGTACCTGTGAGATTGTAGTGGATTGAAATCAGGTGACATCTGGAAG 304

Qy 317 ATACCAGCACCAGTGGAAACAGTGAATTAACAAGCTGAAGCTTTGTTAAGAGCAGCATGCC 376
Db 305 ATACCAGCACCAGTGGAAACAGTGAATTAACAAGCTGAAGCTTTGTTAAGAGCAGCATGCC 364

```


Db 425 TGGCATACCTCTATGAATCTTTAAGTGAAGAGGAGCATGACACAGAAATCCTTT----- 480
Qy 497 CTAACAAGAAATGTGTTCCATGGACCAAGATACCTTCAGTGCAGGTGCAGGCGAG 556
Db 481 -----GATACCTCAGGTGCAGGTGCAGGCGAG 508
Qy 557 GGGCCGATCCCGGTCCTTCGTCGTCGCCCGCCACTCAGGTGTCCTTGAAGGAACCA 616
Db 509 GGGCCGATCCCGGTCCTTCGTCGTCGCCCGCCACTCAGGTGTCCTTGAAGGAACCA 568
Qy 617 AGCCATCAACATCGAGTCAGACCTCTTCCACAGCCTCGGCCCTCTTCCACGAGCCCT 676
Db 569 AGCCATCAACATCGAGTCAGACCTCTTCCACAGCCTCGGCCCTCTTCCACGAGCCCT 628
Qy 677 CTCCTCGAGGCGAGAGCGTTCCTCAAGTTGCGGTGTCGGGGTGTGTCATCTCCCTTA 736
Db 629 CTCCTCGAGGCGAGAGCGTTCCTCAAGTTGCGGTGTCGGGGTGTGTCATCTCCCTTA 688
Qy 737 AAGGAAGTGTCCCTCTGTGGCAAGTGAAGTCTCCAGCTTTGCTCAGCTCTCCCA 796
Db 689 AAGGAAGTGTCCCTCTGTGGCAAGTGAAGTCTCCAGCTTTGCTCAGCTCTCCCA 748
Qy 797 ACAGAAAGTGTGCTTTCGTCGTTGGAACCCAGGATCAGGAGGATTTGGAGCCCG 856
Db 749 ACAGAAAGTGTGCTTTCGTCGTTGGAACCCAGGATCAGGAGGATTTGGAGCCCG 808
Qy 857 TGAAGAGAAATGAGAGGATGGGACCTTCACTGAACGGGAGTTGTTGTCGCAC 916
Db 809 TGAAGAGAAATGAGAGGATGGGACCTTCACTGAACGGGAGTTGTTGTCGCAC 868
Qy 917 AACCGGCTAGAAATGCCAAACCGTCCACGAGGATCAGAGCAGCGGCTGGGAAGCCAG 976
Db 869 AACCGGCTAGAAATGCCAAACCGTCCACGAGGATCAGAGCAGCGGCTGGGAAGCCAG 928
Qy 977 ACAAGATGGAGGAGAGCTGACATGATCATCTGCCAGGACCTGCTGACGAGCTGGGTGA 1036
Db 929 ACAAGATGGAGGAGAGCTGACATGATCATCTGCCAGGACCTGCTGACGAGCTGGGTGA 988
Qy 1037 GTTTGAGCCCTGCATGACACAGCTTCTGCGGCTTGTACTCGGCTGATGGAGCGCT 1096
Db 989 GTTTGAGCCCTGCATGACACAGCTTCTGCGGCTTGTACTCGGCTGATGGAGCGCT 1048
Qy 1097 CGTCCCTGTGCTTACCTGCGGCTGTCGTCGAGCGGATCTGTAAACCACTCTCA 1156
Db 1049 CGTCCCTGTGCTTACCTGCGGCTGTCGTCGAGCGGATCTGTAAACCACTCTCA 1108
Qy 1157 ACACCTCGTGGAGGATACCTCATCCAGCATCCAGAGAGTTCGAGTGAAGAAGATG 1216
Db 1109 ACACCTCGTGGAGGATACCTCATCCAGCATCCAGAGAGTTCGAGTGAAGAAGATG 1168
Qy 1217 TGCAAGATATGGATGCGAGGAATAAATCACTCAAGATGTCGAGCGCCCAAGTCAGGC 1276
Db 1169 TGCAAGATATGGATGCGAGGAATAAATCACTCAAGATGTCGAGCGCCCAAGTCAGGC 1228
Qy 1277 GGTCTTTTCTGTGAAGAGGAGTTTCAGAGACCTGTCGAGCTGTGAGCTGTGACA 1336
Db 1229 GGTCTTTTCTGTGAAGAGGAGTTTCAGAGACCTGTCGAGCTGTGAGCTGTGACA 1288
Qy 1337 GTGAGTCTCAGACATTCAGGACCAATGTCGTCGTCGCGGAGTGTCTGAGTACAGAA 1396
Db 1289 GTGAGTCTCAGACATTCAGGACCAATGTCGTCGTCGCGGAGTGTCTGAGTACAGAA 1348
Qy 1397 GGCAGGCGGCGAGCTCCCACTGCCAGCACCCGAGGCGGAGCCAGGCCCCACAG 1456
Db 1349 GGCAGGCGGCGAGCTCCCACTGCCAGCACCCGAGGCGGAGCCAGGCCCCACAG 1408
Qy 1457 CCCTGGGGGATGCAACCTCCAGCTCCGTCAGCTGACGACAGCAGTCCAGGATTCGTT 1516
Db 1409 CCCTGGGGGATGCAACCTCCAGCTCCGTCAGCTGACGACAGCAGTCCAGGATTCGTT 1468
Qy 1517 GCCCTCTGCAAGGAGCAGCCCTGTGCACTGCTGCTTCCAGCCCTGCGGAGCGGA 1576

Db 1469 GCCCTCTGCAAGGAAGCAAGCCCTGTGACCTGTGCTTCCAGCCCATGCCCGACCGGA 1528
Qy 1577 GAGCGAGCCGAGCAGGACCCGCTGTGCCCTCAGCAGTGTGGGTCTGCTCAGC 1636
Db 1529 GAGCGAGCCGAGCAGGACCCGCTGTGCCCTCAGCAGTGTGGGTCTGCTCAGC 1588
Qy 1637 CTTTCTCCCACTGTACTGGGGTGCACCCGAGCCGCTGCTACGGTGCCTGCCGCCCT 1696
Db 1589 CTTTCTCCCACTGTACTGGGGTGCACCCGAGCCGCTGCTACGGTGCCTGCCGCCCT 1648
Qy 1697 TTTGTAGCTCAACCTGGGTGACAGTGTCTGACCGGCTGTGAACCAACAGCTACG 1756
Db 1649 TTTGTAGCTCAACCTGGGTGACAGTGTCTGACCGGCTGTGAACCAACAGCTACG 1708
Qy 1757 AGTCAGACATCTCTGACAAATTTACCTGCAACACAGAGTTTGACATGAAAAAATGTTGA 1816
Db 1709 AGTCAGACATCTCTGACAAATTTACCTGCAACACAGAGTTTGACATGAAAAAATGTTGA 1768
Qy 1817 CCGAGAGCTCTGCTGCTCTCAGCGGGAGTGTCTGCTGTCTGATTACAGAGTCAAG 1876
Db 1769 CCGAGAGCTCTGCTGCTCTCAGCGGGAGTGTCTGCTGTCTGATTACAGAGTCAAG 1828
Qy 1877 GAGACCCGCTCTGCTGTACTGCTGTGGCTGCGCAGCTTCCGTCAGCTGACCTATCAGT 1936
Db 1829 GAGACCCGCTCTGCTGTACTGCTGTGGCTGCGCAGCTTCCGTCAGCTGACCTATCAGT 1888
Qy 1937 ATCCGAGAAATCTCTGCTTCCAGTTCGAGTGGCGGTAAATCCCGCTCTGACTGCT 1996
Db 1889 ATCCGAGAAATCTCTGCTTCCAGTTCGAGTGGCGGTAAATCCCGCTCTGACTGCT 1948
Qy 1997 ACTGGGCGCTTAATCGCGCACTCAGGTGAAGCTCACCGCCATCAAAATTCATCAT 2056
Db 1949 ACTGGGCGCTTAATCGCGCACTCAGGTGAAGCTCACCGCCATCAAAATTCATCAT 2008
Qy 2057 TCTGTGAACACAGCAAGTTCAAAAATTAAGCATCAGAGGCCCTGAGCAGCTTTCAGCAC 2116
Db 2009 TCTGTGAACACAGCAAGTTCAAAAATTAAGCATCAGAGGCCCTGAGCAGCTTTCAGCAC 2068
Qy 2117 TGGAGTGAAGAGAGCGTGTGTTTTAAATACAGAGACAGCAGCTCAAGGTGTTTTACA 2176
Db 2069 TGGAGTGAAGAGAGCGTGTGTTTTAAATACAGAGACAGCAGCTCAAGGTGTTTTACA 2128
Qy 2177 GCCCCTCAGAGGAAGGAGCAGCGGTCTCCGACAGGTCTCTGGGGTGACTTCTCTGTGG 2236
Db 2129 GCCCCTCAGAGGAAGGAGCAGCGGTCTCCGACAGGTCTCTGGGGTGACTTCTCTGTGG 2188
Qy 2237 AGCTTTTACCTCTGAGTGAGACCTTCCAGAGCCCGGGGCGCCAGCGCCCTCC 2296
Db 2189 AGCTTTTACCTCTGAGTGAGACCTTCCAGAGCCCGGGGCGCCAGCGCCCTCC 2248
Qy 2297 TGCTGAGCGCTGGCAGCGCTCGTGTGGCATCAGCAGCAGAGACGAGCTTCTGTAA 2356
Db 2249 TGCTGAGCGCTGGCAGCGCTCGTGTGGCATCAGCAGCAGAGACGAGCTTCTGTAA 2308
Qy 2357 CATCGCGCTCCCGCAGAGGCGCAGTTTGTCTTTTGTACATTTTCGAAACTACA 2416
Db 2309 CATCGCGCTCCCGCAGAGGCGCAGTTTGTCTTTTGTACATTTTCGAAACTACA 2368
Qy 2417 GTTAAAGCAGAGTCTGTTTTTCAGAAAAAGTTTCAAGGAGAGCGCAAGTTTATCAAAA 2476
Db 2369 GTTAAAGCAGAGTCTGTTTTTCAGAAAAAGTTTCAAGGAGAGCGCAAGTTTATCAAAA 2428
Qy 2477 ACATTTGTTTCAGGAGAGGAGCATAAGTTTACAGCTTACAGGCGTACACAATATCTG 2536
Db 2429 ACATTTGTTTCAGGAGAGGAGCATAAGTTTACAGCTTACAGGCGTACACAATATCTG 2488
Qy 2537 CTGCTGGGAAAAACACAGCATTTTATCTATTTTTTATTTTAAATAGGTTGGTGTATCT 2596
Db 2489 CTGCTGGGAAAAACACAGCATTTTATCTATTTTTTATTTTAAATAGGTTGGTGTATCT 2548
Qy 2597 TCTAATAAGATTAAATCTCAGAACTGTAGCAGAAAAATATAATTTATATTTTACAAA 2656
Db 2549 TCTAATAAGATTAAATCTCAGAACTGTAGCAGAAAAATATAATTTATATTTTACAAA 2608

1621 GCGTCTGCTGAGCCTTTCTGCCACCTGTACTGGGGCTGACCGCGGCTGCTAC 1680
1681 GGTGCTGCGCCCGCTTTGTGAGCTCAACCTGGGTGACAGTGTCTGACCGCGTGTG 1740
1681 GGTGCTGCGCCCGCTTTGTGAGCTCAACCTGGGTGACAGTGTCTGACCGCGTGTG 1740
1741 AACCAACAGCTACGAGTACAGTACAGTCTCTGAGAAATTTACCTGCAACACAGAGTTTGACA 1800
1741 AACCAACAGCTACGAGTACAGTACAGTCTCTGAGAAATTTACCTGCAACACAGAGTTTGACA 1800
1801 TGGAAAAACATGTTGACCGAGAGCTGTGGCTCTCCAGCGGGAGTGTTCCTGCTGTCT 1860
1801 TGGAAAAACATGTTGACCGAGAGCTGTGGCTCTCCAGCGGGAGTGTTCCTGCTGTCT 1860
1861 GATTACAGAGTACCGGAGACCGTTCCTGTTGTTACTGCTGTGCTGCTGCGAGCTTCCGT 1920
1861 GATTACAGAGTACCGGAGACCGTTCCTGTTGTTACTGCTGTGCTGCTGCGAGCTTCCGT 1920
1921 GAGCTGACCTTATCAGTATCGGCGAGAACATTCCTGCTTCCAGTGTGCGAGTGGCGGTAACA 1980
1921 GAGCTGACCTTATCAGTATCGGCGAGAACATTCCTGCTTCCAGTGTGCGAGTGGCGGTAACA 1980
1981 TCCGCTCTGACTGCTACTGGGCGCGTAACCTGCGGCACTCAGTGAAGCTCACACGCC 2040
1981 TCCGCTCTGACTGCTACTGGGCGCGTAACCTGCGGCACTCAGTGAAGCTCACACGCC 2040
2041 ATGAAATTCATATCATATCTGTGAACAGACAGGTTCAAAACTAAGCATCCAGAGCCCT 2100
2041 ATGAAATTCATATCATATCTGTGAACAGACAGGTTCAAAACTAAGCATCCAGAGCCCT 2100
2101 GAGCAGCTTTCAGCACTGAGAGTGAAGAGAGCGTGTGTTTTAAATACAGAGACAAGCAG 2160
2101 GAGCAGCTTTCAGCACTGAGAGTGAAGAGAGCGTGTGTTTTAAATACAGAGACAAGCAG 2160
2161 TCAAGTGTGTTTTCACAGCCCTCAGGAGAGGAGCGAGGTTTCCGACAGGTCCTCTGG 2220
2161 TCAAGTGTGTTTTCACAGCCCTCAGGAGAGGAGCGAGGTTTCCGACAGGTCCTCTGG 2220
2221 GGTGACTCTTCTGTGAGAGCTTTTACCTCTGAGTGAGACCTTCCCGAGAGCCCGGGG 2280
2221 GGTGACTCTTCTGTGAGAGCTTTTACCTCTGAGTGAGACCTTCCCGAGAGCCCGGGG 2280
2281 CCGCAGCCCGCCCTCCTGTTGAGCGCTGGGAGAGGCTCGTGGTGGCATCAGCAGCAGAGA 2340
2281 CCGCAGCCCGCCCTCCTGTTGAGCGCTGGGAGAGGCTCGTGGTGGCATCAGCAGCAGAGA 2340
2341 CGAAGCCTTCTGTAAACATGCGCCCGTCCCGCCGAGAGGGGCAAGTTTGTCTTTGTAC 2400
2341 CGAAGCCTTCTGTAAACATGCGCCCGTCCCGCCGAGAGGGGCAAGTTTGTCTTTGTAC 2400
2401 ATTTTCGAAACTACAGTTAAGCAGAGTGTGTTTTCAGGAAAGTTTTCAGGGAGAAG 2460
2401 ATTTTCGAAACTACAGTTAAGCAGAGTGTGTTTTCAGGAAAGTTTTCAGGGAGAAG 2460
2461 GGCAAGTTTATCAAAAACATTTTTCAGGAGAGGGAGCATAGTTTTCAGCCCTACAGGA 2520
2461 GGCAAGTTTATCAAAAACATTTTTCAGGAGAGGGAGCATAGTTTTCAGCCCTACAGGA 2520
2521 CGTACAAATATCTGCTGCTGGGAAACCAACAGATTTTATCTATTTTATTTTAAATA 2580
2521 CGTACAAATATCTGCTGCTGGGAAACCAACAGATTTTATCTATTTTATTTTAAATA 2580
2581 GGTGTTGCTGCTTATCTTCTAATAGATTTAATGTCACAACTGTAGCAAAATAATATA 2640
2581 GGTGTTGCTGCTTATCTTCTAATAGATTTAATGTCACAACTGTAGCAAAATAATATA 2640
2641 ATTTATATTTTCAAAATTTGACAAAAAATAAAAAA 2679
2641 ATTTATATTTTCAAAATTTGACAAAAAATAAAAAA 2679

RESULT 2

AX877194
LOCUS 2639 bp DNA linear PAT 17-DEC-2003
DEFINITION 12099 from Patent EP1074617.
ACCESSION AX877194
VERSION AX877194.1 GI:40031930
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primers for synthesising full-length cDNA and their use
Patent: EP 1074617-A 12099 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
Location/Qualifiers
1..2639
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
79..2037
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89614.1"
/translation="MESPEEGKQSPPOPWGRLRLGAEAGEPHVLLRKREWTIGRR
GCDLSPFNKLVSGDHCEIYVDEKSGQVLEDTSTVINKLVKKQCTPLOTQDV
IYLVRKNEPHNVAILYESERKGMQESFDSGAGAGADPRVPPSPATQVCF
EPEPSTSLDFTPASSTEPFAGHRSRSCSGGGGSPKSGSVASDEVSS
ASALPDRKTASFSLEPQDLEPVKKMRGDLGLNGQLLAQPRNACTVHEDV
RAAGKPKMDETLTCTIIQDLDHCVSLQPCMTFCAACYSGWMSRLCTCRCPV
BRIKKNHLDNVEAYLIQHPDKRSBEDVDMARKITQDMLQPKVRRFSDBEGS
SEDLLELDNVDSESDISQPVVCRQCEYRROAQPHPCEPGEPCAPCALQPCFL
TSVSTTAVQIVYVPLQSHALCTCCFPMPDRAREEDPRVAPQCAVCLQPCFL
YNGTTRTCYGLAPFCBLNLGDCLDGLNNNSYESDLKNYLAIRGLTKWKLMTES
LVALQGVFLLSRVATGDTVLVCCGGRSFRLETYOYRQNIASELPVAVTSRPDCY
WGRNCRTOGAHAMKFNHICEOTRFKN"

ORIGIN

Query Match 95.8%; Score 2566.4; DB 6; Length 2639;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2613; Conservative 0; Mismatches 1; Indels 36; Gaps 1;
QY 17 GCGCAATCTCTCTTTCAGACGCGCGCGCGCGTTCGGGTTTCGGGCGCGGCGG 76
DB 5 GCGCAATCTCTCTTTCAGACGCGCGCGCGCGTTCGGGTTTCGGGCGCGGCGG 64
QY 77 GGAATGTGAATCCGATGGAGCGCGCGCGCGCGAGGAGGAGTCCGCGCGCGCGCCT 136
DB 65 GGAATGTGAATCCGATGGAGCGCGCGCGCGAGGAGGAGGAGTCCGCGCGCGCCT 124
QY 137 GGGGACGCTCTCTGCTCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 196
DB 125 GGGGACGCTCTCTGCTCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 184
QY 197 GGGAGTGGACCATCGCGCGGAGACGAGGTTGCGACTTTCTTCTTCCCGAGCAATAACTGG 256
DB 185 GGGAGTGGACCATCGCGCGGAGACGAGGTTGCGACTTTCTTCTTCCCGAGCAATAACTGG 244
QY 257 TCTCTGAGATCACTGTAGATTTAGTGGATGAAATTCAGTTCAGTTCAGTTCAGTTCAGT 316
DB 245 TCTCTGAGATCACTGTAGATTTAGTGGATGAAATTCAGTTCAGTTCAGTTCAGTTCAGT 304
QY 317 ATACGACGACCATGAGGAAACAGTGAATTAACAAGCTGAAGGTTCTTTAAGAGCAGACATGCC 376
DB 305 ATACGACGACCATGAGGAAACAGTGAATTAACAAGCTGAAGGTTCTTTAAGAGCAGACATGCC 364
QY 377 CTTTACAGACTGGGATGTCTACTTGGTGTACAGGAGATGACCGGACACCAACG 436
DB 365 CTTTACAGACTGGGATGTCTACTTGGTGTACAGGAGATGACCGGACACCAACG 424
QY 437 TGGCATACTCTATGAATCTTTAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 496

541	Qy	GCAGTGCACGGCGAGGGGCGCATCCC GGGTCTCCTCGTCTGTCGCGCCACTCAGGTG	600
541	Db	GCAGTGCACGGCGAGGGGCGCATCCC GGGTCTCCTCGTCTGCGCGCCACTCAGGTG	600
601	Qy	TGCTTTTCAGGAACACAGGCCATCAACATCGACGTGACCTCTTCCACAGCCTTCGGCC	660
601	Db	TGCTTTTCAGGAACACAGGCCATCAACATCGACGTGACCTCTTCCACAGCCTTCGGCC	660
661	Qy	TCCTTCACGGAGCCTTCTCTGTCAGGGCGAGAGCGTTCTCCAGTTGTGGGTCTCGGGGT	720
661	Db	TCCTTCACGGAGCCTTCTCTGTCAGGGCGAGAGCGTTCTCCAGTTGTGGGTCTCGGGGT	720
721	Qy	GGTGGCATCTCCCTAAAGAAAGTGGTCCCTCTGTGGCAAGTGATGAAGTCTCCAGCTTT	780
721	Db	GGTGGCATCTCCCTAAAGAAAGTGGTCCCTCTGTGGCAAGTGATGAAGTCTCCAGCTTT	780
781	Qy	GCCTCAGCTCTCCGACAGAAAGACTGGTCTCTTTTCTGTGGAAACCCACAGATCAG	840
781	Db	GCCTCAGCTCTCCGACAGAAAGACTGGTCTCTTTTCTGTGGAAACCCACAGATCAG	840
841	Qy	GAGGATTTGGAGCCGCTGAAGAGAAATGAGAGGAGATGGGGA	900
841	Db	GAGGATTTGGAGCCGCTGAAGAGAAATGAGAGGAGATGGGGA	900
901	Qy	CAGTTGTTGGTCGCACAACCGCTAGAAATGCCCAAACCGTCCACGAGGACCTCAGAGCA	960
901	Db	CAGTTGTTGGTCGCACAACCGCTAGAAATGCCCAAACCGTCCACGAGGACCTCAGAGCA	960
961	Qy	CGCGCTGGGAAGCCAGACAAGATGGAGGAGACGCTGACATGCATCATCTGCCAGGACCTG	1020
961	Db	CGCGCTGGGAAGCCAGACAAGATGGAGGAGACGCTGACATGCATCATCTGCCAGGACCTG	1020
1021	Qy	CTGCACGACTGCGTGAGTTTTCAGCCCTGCATGCACACCGTTCTGCGCGGCTTGTCTACTCG	1080
1021	Db	CTGCACGACTGCGTGAGTTTTCAGCCCTGCATGCACACCGTTCTGCGCGGCTTGTCTACTCG	1080

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 11:30:48 ; Search time 16426 Seconds
(without alignments)
7069.036 Million cell updates/sec

Title: US-10-048-046-1
Perfect score: 2679
Sequence: 1 aagaattcgacagagcg.....acaaaaaaaaaaaaaaaa 2679

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_cm.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_nan.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	2679	100.0	2679	9	AF170724	AF170724 Homo sapi
2	2566.4	95.8	2639	6	AX877194	AX877194 Sequence
3	2566.4	95.8	2639	6	BD156534	BD156534 Primer fo
4	2566.4	95.8	2639	9	AK027887	AK027887 Homo sapi
5	2497.2	93.2	3189	9	BC012072	BC012072 Homo sapi
6	2262.4	84.4	3138	6	AX877222	AX877222 Sequence
7	2262.4	84.4	3138	6	BD156548	BD156548 Primer fo
8	2262.4	84.4	3138	9	AK001658	AK001658 Homo sapi
9	2244.6	83.8	3181	6	AX405642	AX405642 Sequence
10	1919.4	71.6	2448	6	AX834949	AX834949 Sequence
11	1919.4	71.6	2448	9	AK097671	AK097671 Homo sapi
12	1500.4	56.0	2297	9	HS0802304	AL037561 Homo sapi
13	1491	55.7	3194	10	BC049792	BC049792 Mus muscu
14	1279.8	47.8	6235	6	BD183291	BD183291 Novel gen
15	961.8	35.9	2257	9	AK090948	AK090948 Homo sapi
16	876.2	32.7	2186	6	AX713476	AX713476 Sequence
17	876.2	32.7	2186	9	AK054917	AK054917 Homo sapi
18	727	27.1	816	6	AX868758	AX868758 Sequence
19	727	27.1	816	6	BD148820	BD148820 Primer fo
20	608.4	22.7	181012	9	AC127070	AC127070 Homo sapi
21	608.4	22.7	181438	2	AC023047	AC023047 Homo sapi
22	583	21.8	824	6	AX868776	AX868776 Sequence
23	583	21.8	824	6	BD148838	BD148838 Primer fo
24	488.8	18.2	518	6	AX873789	AX873789 Sequence
25	488.8	18.2	518	6	BD153851	BD153851 Primer fo
26	440.8	16.5	151210	2	AC139591	AC139591 Rattus no
27	440.8	16.5	273812	2	AC106441	AC106441 Rattus no
28	430.4	16.1	236751	2	AC128407	AC128407 Rattus no
29	416	15.5	61779	2	AC126301	AC126301 Rattus no
30	218	8.1	181438	2	AC023047	AC023047 Homo sapi
31	142.4	5.3	202792	2	AC125898	AC125898 Rattus no
32	139.2	5.2	158897	2	AC118260	AC118260 Mus muscu
33	93.4	3.5	173911	9	AC092490	AC092490 Homo sapi
34	85.2	3.2	338116	2	AC087159	AC087159 Mus muscu
35	82.8	3.1	321708	2	AC087142	AC087142 Mus muscu
36	82.2	3.1	617	6	AX385952	AX385952 Sequence
37	69.4	2.6	601	11	BV002855	BV002855 S208P6639
38	64.4	2.4	165089	9	AC021646	AC021646 Homo sapi
39	55.6	2.1	125020	9	AF429315	AF429315 Homo sapi
40	55.4	2.1	14923	3	AE001394	AE001394 Plasmodi
41	54.8	2.0	125020	9	AF429315	AF429315 Homo sapi
42	53.4	2.0	61052	2	AC123513	AC123513 Dictyoste
43	53.4	2.0	136240	3	AC117070	AC117070 Dictyoste
44	52.6	2.0	110000	3	AC116957_2	Continuation (3 of
45	52.4	2.0	2000	6	AX655393	AX655393 Sequence

ALIGNMENTS

RESULT 1
AF170724
LOCUS AF170724 2679 bp mRNA linear PRI 03-AUG-2000
DEFINITION Homo sapiens cell cycle checkpoint protein CHFR mRNA, complete cds.
ACCESSION AF170724
VERSION AF170724.1 GI:9651169
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2679)
AUTHORS Scolnick,D.M. and Halazonetis,T.D.
TITLE Chfr defines a mitotic stress checkpoint that delays entry into metaphase

This Page Blank (uspto)

Query Match 8.7%; Score 82.5; DB 5; Length 1023;
 Best Local Similarity 29.7%; Pred. No. 4.1;
 Matches 22; Conservative 8; Mismatches 27; Indels 17; Gaps 1;
 QY 31 APOQCAVCLQPFCHLYMGCTGTCYGCCLAPFCE-----LNLGDKCLDG 73
 DB 950 AFNSCMGCKQPFLLHVMKARVCHCCGKIFCESCTSHKPIKRYINTFVRVCDKCFND 1009
 QY 74 VLNNNSYESDILKN 87
 DB 1010 LQSNPSSNSFLND 1023

RESULT 15

Q810G8 PRELIMINARY; PRT; 523 RA.
 AC Q810G8.
 DT 01-WAR-2003 (TRENBLrel. 23, Created)
 DT 01-WAR-2003 (TRENBLrel. 23, Last sequence update)
 DE 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE CG32018-PG (Zyx102.6 isoform).
 GN CG32018 OR ZYX102EF.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazei R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abrill J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dunagan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
 RA Jaisli M., Kalush F., Karpen G.H., Ka Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.B., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kimons I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,

RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Buesam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Dou P.L.E., Doyle C., Dreanek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzales M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwan C., Jalali M., Kruse D., Li P., Mattel B., Moshrefi A.,
 RA Mcintosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirska R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
 RT "Sequencing of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RP Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Seale S.M., Smith E., Shu S., Smutniak P., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
 RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RP Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A.
 RP Flybase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A.
 RP Stonach B.E., Siegrist S., Renfranz P.J., Macalma T., Beckerle M.C.;
 RT "The Zyx102 gene of Drosophila melanogaster."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003846; AA006571.1; -
 DR EMBL; AF219947; AA085868.1; -
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 3.
 DR ProDom; PD000094; LIM; 3.
 DR SMART; SM00132; LIM; 3.
 DR PROSITE; PS00478; LIM DOMAIN_1; 2.
 DR PROSITE; PS00223; LIM DOMAIN_2; 1.
 SQ SEQUENCE 523 AA; 5908 MW; 79BE2A099F9D7890 CRC64;
 Query Match 8.6%; Score 81.5; DB 5; Length 523;
 Best Local Similarity 24.9%; Pred. No. 2.4;
 Matches 44; Conservative 17; Mismatches 63; Indels 53; Gaps 10;
 QY 7 HALTCC-----FQPNP-----DRAEREDPRVAQQCAVCLQPFCHLYMGCT-----RT 52
 DB 348 HIFCTTDCQINLQGGKFFVALDGKPYCEYDYLQTLKSCVCEPILEILRATKPYHP 407
 QY 53 GCYGCCLAPFCELNLDGKCLDGLV-----NNNSYESDILKNYLATRLGTLWKNL----- 100
 DB 408 QCFTCVV--C-----GKSLDGLFTVDATNQNYCIITDFHKKE-APRCVCVKQPIMPDPGQ 459
 QY 101 --TSSLVALOR-----GVLLSDYRYTG-----DTVLVCYCCGLRSFREL 138
 DB 460 EETIRVALDRSFLECYKCEDCGLLLSSEAGRGYPLDDHVLKSCNAKRVQALT 516

Search completed: May 7, 2004, 14:49:06
 Job time : 26.0148 secs

```
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 7.
DR Pfam; PF02494; HVR; 1.
DR Pfam; PF00057; ldl_recept_a; 1.
DR Pfam; PF00059; lectin_C; 1.
DR Pfam; PF00064; sushi; 3.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SW00032; CCP; 3.
DR SMART; SW00034; CLECT; 1.
DR SMART; SW00181; EGF; 13.
DR SMART; SW00192; LDLA; 1.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00041; C TYPE LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00068; LDLA_2; 1.
KW EGF-like domain.
SQ SEQUENCE 2972 AA; 329457 MW; 87D7BA80562F4C74 CRC64;

Query Match 8.9%; Score 83.5; DB 5; Length 2972;
Best Local Similarity 27.1%; Pred. No. 11;
Matches 55; Conservative 12; Mismatches 51; Indels 85; Gaps 15;

QY 10 CTCCTCP--MPDRAERQDPRAVAPQCAVCLQ-----P 41
Db 1407 CLCNPPCPSTELLQR--DQIGDCTLNSCQNGVCLQWGXHDCACNEYFIVDRENGTP 1464

QY 42 FCH-----LWGCRTTCYGC-----LAPCELN-LGDKC-LDGVLNNSYE---SD 83
Db 1465 FCKPNHCLFTNQRNSGYDCGTRENAFLCP--NLNEFGDPCQYEGQLYNNSYTILSA 1522

QY 84 ILKNYLATRLGTLWKNMTELSVALQGVFLLSDYRTGDTVLVCYCGGLRSFRELTYQYRQ 143
Db 1523 GIENATATN-----LCESLDLT-----YAVPN-----TFCV-----Q 1549

QY 144 NIPASELPVAVTSRPPCYWGRNC 166
Db 1550 N-PTSTTP-SIHRCDPCYGCENC 1570

RESULT 13
ID 016004 PRELIMINARY; PRT; 2531 AA.
AC 016004;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Notch homolog.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinzoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Tokopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7654;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=97454256; PubMed=9310331;
RA Sherwood D.R., McClay D.R.;
RT "Identification and localization of a sea urchin Notch homologue:
RT insights into vegetal plate regionalization and Notch receptor
RT regulation."
RL Development 124:3363-3374 (1997).
DR EMBL; AF000634; AA882088.1; -.
DR InterPro; IPR001304; GRAM dom.
DR InterPro; IPR001306; Znf_FVE.
DR Pfam; PF01363; FVE; 1.
DR Pfam; PF03033; Glyco_transf_28; 1.
DR Pfam; PF02893; GRAM; 2.
DR SMART; SM00064; FVE; 1.
DR SMART; SM00568; GRAM; 1.
DR PROSITE; PS0178; ZF_FVE; 1.
DR PROSITE; PS0178; ZF_FVE; 1.
KW Glycosyltransferase; Transferase.
DR InterPro; IPR001438; EGF_II.
SQ SEQUENCE 1023 AA; 114366 MW; 0E3C8C556C5D3709 CRC64;

DR InterPro; IPR006209; EGF like.
DR InterPro; IPR008297; Notch_dom.
DR InterPro; IPR008000; Notch_dom.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 34.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PR00010; EGFBLCOO.
DR PRINTS; PR01452; NOTCH.
DR SMART; SM00248; ANK; 5.
DR SMART; SM00179; EGF_CA; 23.
DR SMART; SM00004; NL; 3.
DR PROSITE; PS00088; ANK_REPEAT; 5.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00010; ASX HYDROXYL; 21.
DR PROSITE; PS00022; EGF_1; 33.
DR PROSITE; PS01186; EGF_2; 25.
DR PROSITE; PS01187; EGF_CA; 20.
DR PIRSF; PIRSF002279; NOTCH; 1.
KW ANK repeat; EGF-like domain; Repeat.
SQ SEQUENCE 2531 AA; 273982 MW; 5BF42BEC627CA303 CRC64;

Query Match 8.8%; Score 83; DB 5; Length 2531;
Best Local Similarity 28.6%; Pred. No. 11;
Matches 30; Conservative 9; Mismatches 40; Indels 26; Gaps 6;

QY 5 GSHALCTCCPOP--MPDRAERQDPRAVAPQCAVCLQPECHLYWGCTR-TGCVGCL--- 58
Db 840 GYSCRTSGFGNFCDDRNE-----CLFSFGRNGSGCTNLEGSFECCLP 885

QY 59 ---APFCELNLDGKLDGVLNNSYESDILKNYLAT--RGLTWKN 98
Db 886 GYDGPICEINI-DECSGPGCTNGICITLDDIFCSCQRGFTGKN 929

RESULT 14
Q9XYD4 PRELIMINARY; PRT; 1023 AA.
ID Q9XYD4
AC Q9XYD4;
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Sterol glycosyltransferase (EC 2.4.1.173).
GN UGT52.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
[1]
RN SEQUENCE FROM N.A.
RX STRAIN=AX4;
RX MEDLINE=99240683; PubMed=10224056;
RA Warnecke D., Erdmann R., Fahl A., Hube B., Muller F., Zank T.,
RA Zahring U., Heinz E.;
RT "Cloning and functional expression of UGT genes encoding sterol
RT glucosyltransferases from Saccharomyces cerevisiae, Candida albicans,
RT Pichia pastoris, and Dictyostelium discoideum."
RL J. Biol. Chem. 274:13048-13059 (1999).
DR EMBL; AF098916; AAD28546.1; -.
DR GO; GO:0016906; P:sterol 3-beta-glucosyltransferase activity; IEA.
DR GO; GO:0016758; P:transferase activity; transferring hexosyl . . .; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0030259; P:lipid glycosylation; IEA.
DR InterPro; IPR004276; Glyco_trans_28.
DR InterPro; IPR004182; GRAM dom.
DR InterPro; IPR00306; Znf_FVE.
DR Pfam; PF01363; FVE; 1.
DR Pfam; PF03033; Glyco_transf_28; 1.
DR Pfam; PF02893; GRAM; 2.
DR SMART; SM00064; FVE; 1.
DR SMART; SM00568; GRAM; 1.
DR PROSITE; PS0178; ZF_FVE; 1.
DR PROSITE; PS0178; ZF_FVE; 1.
KW Glycosyltransferase; Transferase.
DR InterPro; IPR001438; EGF_II.
SQ SEQUENCE 1023 AA; 114366 MW; 0E3C8C556C5D3709 CRC64;
```


DR	InterPro; IPR001870; B302.	
DR	InterPro; IPR006574; PRY.	
DR	InterPro; IPR003877; SPRY receptor.	
DR	InterPro; IPR000315; Znf_Box.	
DR	InterPro; IPR001841; Znf_Box.	
DR	Pfam; PF00622; SPRY; 1.	
DR	Pfam; PF00643; zf-B_box; 1.	
DR	Pfam; PF00697; zf-C3HC4; 1.	
DR	SMART; SM00336; BBOX; 1.	
DR	SMART; SM00589; PRY; 1.	
DR	SMART; SM00184; RING; 1.	
DR	PROSITE; PS00119; ZF_BBOX; 1.	
DR	PROSITE; PS00518; ZF_RING_1; 1.	
DR	PROSITE; PS00589; ZF_RING_2; 1.	
KW	Hypochemical protein; Metal-binding; Zinc; Zinc-finger.	
FT	NON_TER 1	
SEQ	SEQUENCE 504 AA; 57735 MW; E0B377F32A36D1B9 CRC64;	
	Query Match 9.0%; Score 85; DB 4; Length 504;	
	Best Local Similarity 24.8%; Pred.No. 0.91;	
	Matches 34; Conservative 15; Mismatches 58; Indels 30; Gaps 7;	
QY	16 EMDRRA-BREODPRVAPOQ-----CAVCIQPPCH-LYWGCTRTGCYGCCLAPPCELN 66	
DB	5 PEPAARAMERSPDVSPGSRSPKEILLCAVCYDPFDDAVTLRGCHNFCRGCVSRCEVQV 64	
QY	67 GDKCLGDLVNNYSYEDILKNYL-----ATRGLTWKMLTESLVALQRG---VF 112	
DB	65 SPTC--FVCKDRASPADLTNTLTNLVLEKLLREEAGARWTSYRFSRVCLRHGGQLSF 122	
QY	113 LLSDYRVVTGDTVLCYCC 129	
DB	123 CLED----KELLCSC 134	
RESULT 12		
IP00891	PRELIMINARY; PRT; 2972 AA.	
ID	P00891	
AC	P00891;	
DT	01-MAY-1997 (TrEMBLrel. 03, Created)	
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	F55H12.3 protein.	
DE	F55H12.3.	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	
OC	Rhabditidae; Peloderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Dobson R.;	
RL	Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=99069613; PubMed=9851916;	
RA	none;	
RT	'genome sequence of the nematode C.elegans: A platform for	
RT	investigating biology.';	
RL	Science 282:2012-2018(1998).	
DR	EMBL; Z81091; CAB03143.2; -.	
DR	PIR; T22759; T22759.	
DR	HSSP; P00740; LEDM.	
DR	WormFep; F55H12.3; CF25008.	
DR	GO; GO:000509; F:calcium ion binding; IEA.	
DR	GO; GO:0005198; F:structural molecule activity; IEA.	
DR	GO; GO:0005529; F:sugar binding; IEA.	
DR	InterPro; IPR000152; Asx_hydroxyl_S.	
DR	InterPro; IPR001881; EGF_Ca	
DR	InterPro; IPR006209; EGF-like.	
DR	InterPro; IPR003410; HyalIn.	
DR	InterPro; IPR006210; IEGF.	
DR	InterPro; IPR002049; Laminin EGF.	
DR	InterPro; IPR002172; LDL receptor A.	

Query Match 95.5%; Score 901; DB 11; Length 663;
 Best Local Similarity 93.4%; Pred. No. 3.2e-92;
 Matches 155; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 1 CPLQSHALCTCCFQMPDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 60
 DB 475 CPLQSHALCTCCFQMPDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 534

QY 61 FCELNGLDKCLDGLVNNNYESDILKNYLATRLGTLWKNMLESVALQGVFLLSDYRT 120
 DB 535 FCELNGLDKCLDGLVNNNYESDILKNYLATRLGTLWKNMLESVALQGVFLLSDYRT 594

QY 121 GDTVLYCCGLRSFRLTYQYQNIIPASELPVAVTSRPPCYGGRNC 166
 DB 595 GNTVLYCCGLRSFRLTYQYQNIIPASELPVAVTSRPPCYGGRNC 640

RESULT 8
 Q9SX88 PRELIMINARY; PRT; 473 AA.
 AC Q9SX88;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE F16N3.15.
 GN F16N3.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
 RA Lee J., Liu A., Li J., Kremenetskaia I., Gonzalez A., Altafi H.,
 RA Araujo R., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,
 RA Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C., Rowley D.,
 RA Shinn P., Walker M., Davis R.W., Ecker J.R., Federspiel N.A.,
 RA Theologis A.; thaliana chromosome 1 BAC F16N3 sequence."
 RT "Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases."
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL; AC007519; AAD46042.1; -
 DR PIR; C96516;
 DR InterPro; IPR001841; Znf ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS50089; ZF_RING_2; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 473 AA; 53874 MW; D17217493A81C991 CRC64;

Query Match 14.6%; Score 137.5; DB 10; Length 473;
 Best Local Similarity 25.0%; Pred. No. 1.1e-06;
 Matches 45; Conservative 26; Mismatches 88; Indels 21; Gaps 6;

QY 1 CPLQSHALCTCCFQMPDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 60
 DB 275 CEHGAHLQCHLCQGMPPFRANLQ-----VPLHCKGCDRPFQAYWS-SENVTQVSGP 327

QY 61 FCELN---LGDKLDGV-----LNNNSYESDILKNYLATRLGTLWKNMLES-VAL 107
 DB 328 VCVRETRFISERTITRPFITHEMRHQDITQRCIAHEKTPDVVAELWLFNREI 387

QY 108 QRGVFLLSDYR-VGTDTVLYCCGLRSFRLTYQYQNIIPASELPVAVTSRPPCYGGRNC 166
 DB 388 DRSRMLNHAETITASTHVNCDCYDKLVGFLYWFRLTLPRLNHLPAVAREDCWGVAC 447

RESULT 9
 Q7T3F2 PRELIMINARY; PRT; 292 AA.
 ID Q7T3F2

AC Q7T3F2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalusz D.E., Scherch A., Schein J.B.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC053142; AAHS3142.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 292 AA; 31678 MW; 5BBA8E434A66287C CRC64;

Query Match 9.7%; Score 91.5; DB 13; Length 292;
 Best Local Similarity 23.2%; Pred. No. 0.089;
 Matches 35; Conservative 20; Mismatches 67; Indels 29; Gaps 7;

QY 1 CPLQSHALCTCCFQMPDRAEREDPRVAP-----OQCAVCLQPFCHLY---- 46
 DB 125 CPVCQT-VFCCCRSPNSDGHSCSLQQLPMPPLPDGALSDDTEAALKQCPVCGVIERN 183

QY 47 WGTTRTCYGCCLAPFCELNGLDKCLQGVNNNYESDILKNYL--ATRGLTWN----- 98
 DB 184 QCAQMLCKSKCHTFCWVCLQN--LDGDIPLRHYDKGPCENKLGHSRASVWNRVTQVGI 241

QY 99 MLTSLVALQGVFLLSDYRTGDTVLYCYCC 129
 DB 242 LVGASIIIVLTSPLL-----LASPCILCCVC 268

RESULT 10
 Q9N9P8 PRELIMINARY; PRT; 806 AA.
 ID Q9N9P8
 AC Q9N9P8;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN L2802.02.
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]

Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPLQSHALCTCCQPMDDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 60
DB 435 CPLQSHALCTCCQPMDDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 494

QY 61 FCELNIGDKLDGVLNNNSYSDILKNYLATRGLTWKMLTESLVALQRGVFLLSDYRVT 120
DB 495 FCELNIGDKLDGVLNNNSYSDILKNYLATRGLTWKMLTESLVALQRGVFLLSDYRVT 554

QY 121 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGNC 166
DB 555 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGNC 600

RESULT 5

Q96EP1 PRELIMINARY; PRT; 652 AA.

AC Q96EP1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Placenta;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; BC012072; AAH12072.1; -;
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; ZF-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS0006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 652 AA; 72058 MW; 1A50773CEA679F07 CRC64;

Query Match 99.6%; Score 939; DB 4; Length 652;
Best Local Similarity 99.4%; Pred. No. 1.7e-96;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CPLQSHALCTCCQPMDDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 60
DB 464 CPLQSHALCTCCQPMDDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 523

QY 61 FCELNIGDKLDGVLNNNSYSDILKNYLATRGLTWKMLTESLVALQRGVFLLSDYRVT 120
DB 524 FCELNIGDKLDGVLNNNSYSDILKNYLATRGLTWKMLTESLVALQRGVFLLSDYRVT 583

QY 121 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGNC 166
DB 584 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGNC 629

RESULT 6

Q810L3 PRELIMINARY; PRT; 664 AA.

AC Q810L3;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE RIKEN cDNA 5730484M20 Gene.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Limb;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049792; AAH49792.1; -;
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; ZF-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 2.
DR PROSITE; PS0006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 664 AA; 73871 MW; D651BE3E463DEBB6 CRC64;

Query Match 96.5%; Score 910; DB 11; Length 664;
Best Local Similarity 94.0%; Pred. No. 3.1e-93;
Matches 156; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 CPLQSHALCTCCQPMDDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 60
DB 476 CPLQSHALCTCCQPMDDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 535

QY 61 FCELNIGDKLDGVLNNNSYSDILKNYLATRGLTWKMLTESLVALQRGVFLLSDYRVT 120
DB 536 FCELNIGDKLDGVLNNNSYSDILKNYLATRGLTWKMLTESLVALQRGVFLLSDYRVT 595

QY 121 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGNC 166
DB 596 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGNC 641

RESULT 7

Q8BJZ9 PRELIMINARY; PRT; 663 AA.

AC Q8BJZ9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to cell cycle checkpoint protein CHFR.
GN 5730484M20RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK077629; BAC36912.1; -;
DR MGD; MGI:2444898; 5730484M20RIK.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; ZF-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 2.
DR PROSITE; PS0006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 663 AA; 73855 MW; 42BD653DCB4ADF45 CRC64;

```

RESULT 2
Q96SL3 PRELIMINARY; PRT; 652 AA.
AC Q96SL3 2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ14781.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki Y., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK027687; BAB55297.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 652 AA; 72030 MW; 572F2CE6D1743D80 CRC64;

Query Match 100.0%; Score 943; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 6e-97; Indels 0; Gaps 0;
Matches 166; Conservative 0; Mismatches 0;

QY 1 CPLOGSHALCTCCFQMPDRRAERQDPVAPQCAVCLQFFCHLYWGCTRTGCGCLAP 60
DB 464 CPLOGSHALCTCCFQMPDRRAERQDPVAPQCAVCLQFFCHLYWGCTRTGCGCLAP 523
QY 61 FCEINLGDKCLDGVNNSYESDILKNYLATRGITWKNMLTESLVALQGVFLLSDYRVT 120
DB 524 FCEINLGDKCLDGVNNSYESDILKNYLATRGITWKNMLTESLVALQGVFLLSDYRVT 593
QY 121 GDTVLCYCCGLRSFRELTYQYRONIPASELPVATSRPDCYWGNC 166
DB 584 GDTVLCYCCGLRSFRELTYQYRONIPASELPVATSRPDCYWGNC 629

RESULT 3
Q9NRT4 PRELIMINARY; PRT; 664 AA.
AC Q9NRT4 2001 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Cell cycle checkpoint protein CHFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki Y., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK001658; BAA91817.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 664 AA; 73386 MW; 141A1E7FEFAE36A2 CRC64;

Query Match 100.0%; Score 943; DB 4; Length 664;
Best Local Similarity 100.0%; Pred. No. 6.1e-97; Indels 0; Gaps 0;
Matches 166; Conservative 0; Mismatches 0;

QY 1 CPLOGSHALCTCCFQMPDRRAERQDPVAPQCAVCLQFFCHLYWGCTRTGCGCLAP 60
DB 476 CPLOGSHALCTCCFQMPDRRAERQDPVAPQCAVCLQFFCHLYWGCTRTGCGCLAP 535
QY 61 FCEINLGDKCLDGVNNSYESDILKNYLATRGITWKNMLTESLVALQGVFLLSDYRVT 120
DB 536 FCEINLGDKCLDGVNNSYESDILKNYLATRGITWKNMLTESLVALQGVFLLSDYRVT 595
QY 121 GDTVLCYCCGLRSFRELTYQYRONIPASELPVATSRPDCYWGNC 166
DB 596 GDTVLCYCCGLRSFRELTYQYRONIPASELPVATSRPDCYWGNC 641

RESULT 4
Q9NVD5 PRELIMINARY; PRT; 623 AA.
AC Q9NVD5 2001 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ10796.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki Y., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsumura H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK001658; BAA91817.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 623 AA; 69204 MW; 45186D3DAE52711 CRC64;

Query Match 99.6%; Score 939; DB 4; Length 623;
Best Local Similarity 99.4%; Pred. No. 1.6e-96;

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:40:47 ; Search time 24.0148 Seconds
(without alignments)
2180.991 Million cell updates/sec

Title: US-10-048-046-2_COPY_476_641
Perfect score: 943
Sequence: 1 CPLOGSHALCTCCQPMRDR.....ASELPVAVTSRDPYWGRCN 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25:

- 1: sp_archaea:
- 2: sp_bacteria:
- 3: sp_fungi:
- 4: sp_human:
- 5: sp_invertebrate:
- 6: sp_mammal:
- 7: sp_mhc:
- 8: sp_organelle:
- 9: sp_phage:
- 10: sp_plant:
- 11: sp_rodent:
- 12: sp_virus:
- 13: sp_vertebrate:
- 14: sp_unclassified:
- 15: sp_virus:
- 16: sp_bacteriap:
- 17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	943	100.0	306	4 Q9NT32	Q9NT32 homo sapien
2	943	100.0	652	4 Q96SL3	Q96SL3 homo sapien
3	943	100.0	664	4 Q9NRT4	Q9NRT4 homo sapien
4	939	99.6	623	4 Q9NVD5	Q9NVD5 homo sapien
5	939	99.6	652	4 Q96EP1	Q96EP1 homo sapien
6	910	96.5	664	11 Q810L3	Q810L3 mus musculus
7	901	95.5	663	11 Q8BJZ9	Q8BJZ9 mus musculus
8	137.5	14.6	473	10 Q9SX88	Q9SX88 arabidopsis
9	91.5	9.7	292	13 Q7T3P2	Q7T3P2 brachydanio
10	88.5	9.4	806	5 Q9N9P8	Q9N9P8 leishmania
11	85	9.0	504	4 Q9UPQ4	Q9UPQ4 homo sapien
12	83.5	8.9	2972	5 P90891	P90891 caenorhabdi
13	83	8.8	2531	5 O16004	O16004 lytechinus
14	82.5	8.7	1023	5 Q9XYD4	Q9XYD4 dictyosteli
15	81.5	8.6	523	5 Q810G8	Q810G8 drosophila
16	81.5	8.6	564	5 Q8T0F5	Q8T0F5 drosophila

17	81.5	8.6	585	5 Q9N675	Q9N675 drosophila
18	81.5	8.6	1056	11 Q8COP7	Q8COP7 mus musculus
19	81	8.6	551	2 Q8L172	Q8L172 desulfitoba
20	81	8.6	551	2 Q8GJ31	Q8GJ31 desulfitoba
21	81	8.6	551	2 Q8GJ27	Q8GJ27 dehalobacte
22	80.5	8.5	274	12 Q7TS17	Q7TS17 cryptosporid
23	80.5	8.5	336	2 Q69207	Q69207 actinosyne
24	80	8.5	302	10 Q64762	Q64762 arabidopsis
25	80	8.5	530	5 Q9VTV3	Q9VTV3 drosophila
26	80	8.5	628	13 Q7ZVC1	Q7ZVC1 brachydanio
27	79	8.4	1532	4 Q9P2K1	Q9P2K1 homo sapien
28	78.5	8.3	206	4 Q8WVA4	Q8WVA4 homo sapien
29	78.5	8.3	493	4 Q86XQ0	Q86XQ0 homo sapien
30	78	8.3	310	11 Q8CPU8	Q8CPU8 mus musculus
31	78	8.3	500	11 Q99PN4	Q99PN4 mus musculus
32	78	8.3	639	11 Q8K450	Q8K450 mus musculus
33	77.5	8.2	446	13 Q8Q3L6	Q8Q3L6 brachydanio
34	77.5	8.2	601	13 Q8JH43	Q8JH43 brachydanio
35	77.5	8.2	1060	3 Q8TFW3	Q8TFW3 aspergillus
36	77.5	8.2	2215	5 Q81BW9	Q81BW9 plasmodium
37	77	8.2	303	4 Q7Z419	Q7Z419 homo sapien
38	77	8.2	1633	11 Q8CFW7	Q8CFW7 mus musculus
39	76.5	8.1	1358	5 Q960D5	Q960D5 drosophila
40	76.5	8.1	1515	13 Q9DE37	Q9DE37 brachydanio
41	76.5	8.1	1700	5 Q9VVA9	Q9VVA9 drosophila
42	76.5	8.1	2841	5 Q8MLU9	Q8MLU9 drosophila
43	76.5	8.1	2931	5 Q9W2C6	Q9W2C6 drosophila
44	76	8.1	120	2 Q8GHL2	Q8GHL2 serratia ma
45	76	8.1	181	10 Q9AWV5	Q9AWV5 cryza sativ

ALIGNMENTS

RESULT 1

Q9NT32 PRELIMINARY; PRT; 306 AA.
AC Q9NT32;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DE 01-JUN-2003 (TREMREL. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP434N2420.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Testis;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI37561; CAB70812.1; -
DR PIR; T46399; T46399.
KW Hypothetical protein.

FT NON_TER
SQ SEQUENCE 306 AA; 34501 MW; 6B50F04601FB2939 CRC64;
Query Match 100.0%; Score 943; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.4e-97;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CPLOGSHALCTCCQPMRDRAREODPRVAPQCCAVCIQPFCHLYWGCTRTGCYCCLAP	60
Db	118	CPLOGSHALCTCCQPMRDRAREODPRVAPQCCAVCIQPFCHLYWGCTRTGCYCCLAP	177
Qy	61	FCELNLDGKCLDGLVNNNSVESDILKYLATRLGLTWKNNLTSLVALQRGVLLSDYRT	120
Db	178	FCELNLDGKCLDGLVNNNSVESDILKYLATRLGLTWKNNLTSLVALQRGVLLSDYRT	237
Qy	121	GDTVLVCYCGLSFRELTYQYRONIPASELPVAVTSRDPYWGRCN	166
Db	238	GDTVLVCYCGLSFRELTYQYRONIPASELPVAVTSRDPYWGRCN	283

DR PROSITE; PS50023; LIM DOMAIN 2; 4.
KW Repeat; LIM domain; Metal-binding; Zinc; Zinc-finger.
FT ZN FING 7 31 GATA-LIKE (POTENTIAL).
FT DOMAIN 40 92 LIM 1.
FT DOMAIN 101 153 LIM 2.
FT DOMAIN 162 212 LIM 3.
FT DOMAIN 221 275 LIM 4.
SQ SEQUENCE 279 AA; 32072 MW; 6D8CB4B4424BFF2 CRC64;

Query Match 8.0%; Score 75; DB 1; Length 279;
Best Local Similarity 23.6%; Pred. No. 3.2; Mismatches 12; Indels 34; Gaps 6;
Matches 29; Conservative 12; Mismatches 48; Indels 34; Gaps 6;

QY 9 LCTCCFQMPDRR-AEREOPRVAPOQCAVCLQFFCHLYWGCTRTGCGYCLAPFCELNLG 67
Db 187 VCTACKQLSGQRTARDEFP-----YCLTCFCDLY-----AKCAGCTNPIISGLG-G 233
QY 68 DKCLDGLVNNNSVESDILKYLATRGITWKNMLTESLVALQGVFLSDYRVGTGTVLCY 127
Db 234 TKYIS--FEERQWHDNCFNCKGSLSLVGRGFLTER-----DDILCP 273
QY 128 CCG 130
Db 274 DCG 276

RESULT 14
SLI3 RAT STANDARD; PRT; 279 AA.
AC O35115; 2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL)
DE (Four and a half LIM domains protein 2) (FHL-2).
GN FHL2 OR SLIM3 OR DRAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Wistar; TISSUE=Brain;
RA Tanahashi H.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -/- SIMILARITY: Contains 4 LIM zinc-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB008571; BAA23357.1; -
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 4.
DR ProDom; PD000094; LIM; 4.
DR SMART; SM00132; LIM; 4.
DR PROSITE; PS00478; LIM DOMAIN 1; 4.
DR PROSITE; PS50023; LIM DOMAIN 2; 4.
KW Repeat; LIM domain; Metal-binding; Zinc; Zinc-finger.
FT ZN FING 7 31 GATA-LIKE (POTENTIAL).
FT DOMAIN 40 92 LIM 1.
FT DOMAIN 101 153 LIM 2.
FT DOMAIN 162 212 LIM 3.
FT DOMAIN 221 275 LIM 4.
SQ SEQUENCE 279 AA; 32086 MW; 9A9DBE5935034173 CRC64;

Query Match 8.0%; Score 75; DB 1; Length 279;
Best Local Similarity 23.6%; Pred. No. 3.2;
Matches 29; Conservative 12; Mismatches 48; Indels 34; Gaps 6;

QY 9 LCTCCFQMPDRR-AEREOPRVAPOQCAVCLQFFCHLYWGCTRTGCGYCLAPFCELNLG 67
Db 187 VCTACKQLSGQRTARDEFP-----YCLTCFCDLY-----AKCAGCTNPIISGLG-G 233
QY 68 DKCLDGLVNNNSVESDILKYLATRGITWKNMLTESLVALQGVFLSDYRVGTGTVLCY 127
Db 234 TKYIS--FEERQWHDNCFNCKGSLSLVGRGFLTER-----DDILCP 273
QY 128 CCG 130
Db 274 DCG 276

RESULT 15
OPCA NOSPU STANDARD; PRT; 465 AA.
AC P48971;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative OXPP cycle protein opca.
GN OPCA.
OS Nostoc punctiforme.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=63737;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 29133 / PCC 73102;
RX MEDLINE=95175603; PubMed=7870816;
RA Summers M.L., Meeks J.C., Chu S., Wolf R.E. Jr.;
RT "Nucleotide sequence of an operon in Nostoc sp. strain ATCC 29133
RT encoding four genes of the oxidative pentose phosphate cycle.";
RL Plant Physiol. 107:267-268(1995).
CC -/- FUNCTION: MAY BE INVOLVED IN THE FUNCTIONAL ASSEMBLY OF GLUCOSE
CC 6-PHOSPHATE DEHYDROGENASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32796; AAA50771.1; -
DR InterPro; IPR004555; OPCA.
DR InterPro; IPR002477; PG binding.
DR Pfam; PF01471; PG_binding_1; 1.
DR TIGRFAMs; TIGR00534; OPCA; 1.
SQ SEQUENCE 465 AA; 50627 MW; 2E2D0B0D90E23931 CRC64;

Query Match 8.0%; Score 75; DB 1; Length 465;
Best Local Similarity 22.8%; Pred. No. 5.6;
Matches 38; Conservative 19; Mismatches 68; Indels 42; Gaps 7;

QY 1 CPL--QGSHALCTCCFQMPDRRAREQDPRAVQCAVCLQFFCHLYWGCT---RTGCV 55
Db 183 CPIQKSSSTLTCCEYITLSGTAALERIGGMIPALLIGGLPKF--LWKATPDNNGLF 240
QY 56 GCLAPFCELNLGDKCLDGLVNNNSVESDILK-NYLATRGL-----TWKNMLTES 103
Db 241 KSLAALCNVIVDSC-----NFNEPESDLRLLEELVEAGVPLADLNWRRLASQWELTAZ 295
QY 104 LVALQGVFLSDYRVGTGTVLCYCCGLRSFELTYQYRONIPASEL 150
Db 296 YDSPKRR-----AALREIDRVTDYKGNPAQAL 324

Search completed: May 7, 2004, 14:46:34
Job time : 6.60929 secs

QY 135 RELTYQYRQNP 146
DB 64 BELTGRYRNP 75

RESULT 13

SLI3_MOUSE
ID SLI3_MOUSE STANDARD; PRT; 279 AA.
AC 070433; P97448;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL)
GN (Four and a half LIM domains protein 2) (FHL-2).
GN FHL2.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RA Chan K.K., Tsui S.K.W., Lee C.Y., Fung K.P., Waye M.M.Y.;
RT "The cloning, sequencing and characterization of a mouse FHL2, which
RT contains four and a half LIM domains."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=99160848; PubMed=10049693;
RA Morgan M.J., Madgwick A.J.A.;
RT "The LIM proteins FHL1 and FHL3 are expressed differently in skeletal
RT muscle."
RL Biochem. Biophys. Res. Commun. 255:245-250(1999).
RN [3]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=20368480; PubMed=10906474;
RA Chu P.-H., Ruiz-Lozano P., Zhou Q., Cai C., Chen J.;
RT "Expression patterns of FHL/SLIM family members suggest important
RT functional roles in skeletal muscle and cardiovascular system."
RL Mech. Dev. 95:259-265(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Starzinski-Powitz A., Martin B., Eckardt F.;
RT "Isolation of the mouse homolog mDRAL from skeletal muscle derived
RT myoblasts."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- TISSUE SPECIFICITY: Highly expressed in heart but also detectable
CC in brain and skeletal muscle.
CC -1- SIMILARITY: Contains 4 LIM zinc-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF055889; AAC12770.1; -;
DR EMBL; U77040; AAB19211.2; -;
DR EMBL; AF114381; AAD53230.1; -;
DR EMBL; AF153340; AAD34170.1; -;
DR MGD; MGI:1338762; FHL2.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0006357; P:regulation of transcription from Pol II promoter; IDA.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 4.
DR ProDom; PD000094; LIM; 4.
DR SMART; SM00132; LIM; 4.
DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
DR

RESULT 12

LPCA_YERPE
ID LPCA_YERPE STANDARD; PRT; 193 AA.
AC Q8ZBY7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoglycerate isomerase (EC 5.4.2.1).
GN LPCA OR GWAH OR YPO3243 OR Y0947.
GN Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Irarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moulé S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
CC -1- FUNCTION: Involved in synthesis of glyceromannoheptose 7-phosphate
CC (by similarity).
CC
CC -1- PATHWAY: Inner core lipopolysaccharide biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the SIS family. LPCA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ414156; CAC92478.1; -;
DR EMBL; AE013698; AAM84528.1; -;
DR PIR; AB0394; AB0394.
DR HAMAP; MF_00067; -; 1.
DR InterPro; IPR004515; GnhA.
DR InterPro; IPR001347; SIS.
DR Pfam; PF01380; SIS; 1.
DR TIGRFAMs; TIGR00441; GnhA; 1.
DR Isomerase; Lipopolysaccharide biosynthesis; Complete proteome.
FT CONFLICT 109 109 D -> G (IN REF. 2).
SQ SEQUENCE 193 AA; -20987 MW; 643FEFC034452D79 CRC64;

Query Match 8.0%; Score 75; DB 1; Length 193;
Best Local Similarity 29.2%; Pred. No. 2.1;
Matches 21; Conservative 16; Mismatches 27; Indels 8; Gaps 3;
QY 83 DILKNYATLRLTWKMLTE--SLVALQGVFLSLD-YRVTDVLC-----YCCGLRSF 134
DB 4 DLIRESNEADTLANFLKDSNIDAIQRAAILLADSFKAGGKVLSCNGSGHCDAMHFA 63

[9] CARBOHYDRATE-LINKAGE SITE THR-109.
 MEDLINE=92184750; PubMed=1544894;
 Harris R.J., Ling V.T., Spellman M.W.;
 "O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C";
 J. Biol. Chem. 267:5102-5107 (1992).
 [10]
 VARIANT WASHINGTON D.C. SER-590.
 MEDLINE=90046788; PubMed=2510163;
 Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B., Saito H.;
 "Coagulation factor XII (Hageman factor) Washington D.C.: inactive factor XIIa results from Cys-571-->Ser substitution";
 Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322 (1989).
 [11]
 VARIANT LOCARNO PRO-372.
 MEDLINE=94325559; PubMed=8049433;
 Hovings J.K., Schaller J., Stricker H., Wuillemin W.A., Furlan M., Laemmle B.;
 "Coagulation factor XII Locarno: the functional defect is caused by the amino acid substitution Arg-353-->Pro leading to loss of a kallikrein cleavage site";
 Blood 84:1173-1181 (1994).
 [12]
 VARIANT TENRI CVS-53.
 MEDLINE=92290785; PubMed=10361128;
 Kondo S., Tokunaga F., Kawano S., Ono Y., Kumagai S., Koide T.;
 "Factor XII Tenri, a novel cross-reacting material negative factor XII deficiency, occurs through a proteasome-mediated degradation";
 Blood 93:4300-4308 (1999).
 -!- FUNCTION: Factor XII is a serum glycoprotein that participates in the initiation of blood coagulation, fibrinolysis, and the generation of bradykinin and angiotensin.
 -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-Ile bonds in factor VII to form factor VIIa and factor XI to form factor XIa.
 -!- PTM: O- AND N-GLYCOSYLATED.
 -!- DISEASE: Defects in F12 do not cause any clinical symptoms. The sole effect is that whole-blood clotting time is prolonged.
 -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a complex bound to an anionic surface. Prekallikrein is cleaved by factor XII to form kallikrein, which then cleaves factor XII first to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor XIIa activates factor XI to factor XIa.
 -!- SIMILARITY: Belongs to peptidase family S1.
 -!- SIMILARITY: Contains 2 EGF-like domains.
 -!- SIMILARITY: Contains 1 fibronectin type I domain.
 -!- SIMILARITY: Contains 1 fibronectin type II domain.
 -!- SIMILARITY: Contains 1 kringle domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; M31315; AAA70225.1; -.
 EMBL; AF538691; AAM97932.1; -.
 EMBL; M11723; AAA51986.1; -.
 EMBL; M17466; AAB59490.1; -.
 EMBL; M17464; AAB59490.1; JOINED.
 EMBL; M17465; AAB59490.1; JOINED.
 EMBL; M13147; AAA70224.1; -.
 EMBL; U71274; AAB51203.1; -.
 PIR; A29411; KFHU12.
 HSSP; P00763; 1DPO.
 MESOP; S01-211; -.
 Genew; HGNC:3530; F12.
 MIM; 234000; -.
 GO; GO:0003805; F: blood coagulation factor XI activity; TAS.
 GO; GO:0003806; F: blood coagulation factor XII activity; TAS.

DR GO; GO:0008236; F: serine-type peptidase activity; TAS.
 DR GO; GO:0007596; F: blood coagulation; TAS.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000083; Fibrinctn1.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO00722; CHIMOTRYPsin.
 DR PRINTS; PRO0013; FNTYPEII.
 DR PRINTS; PRO0018; KRINGLE.
 DR PRINTS; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kring1; 1.
 DR SMART; SM00181; EGF_2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00266; EGF_3; 2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Blood coagulation; plasma; Kringle; Serine protease;
 KW Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
 KW Polymorphism; Disease mutation.
 FT SIGNAL 1 19
 FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.
 FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.
 FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.
 FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.
 FT DOMAIN 47 88 FIBRONECTIN TYPE-II.
 FT DOMAIN 94 131 EGF-LIKE 1.
 FT DOMAIN 133 173 FIBRONECTIN TYPE-I.
 FT DOMAIN 174 210 EGF-LIKE 2.
 FT DOMAIN 217 295 KRINGLE.
 FT DOMAIN 296 349 PRO-RICH.
 FT DOMAIN 373 615 SERINE PROTEASE.
 FT CARBOHYD 109 109 O-LINKED (FUC).
 FT CARBOHYD 249 249 O-LINKED (GLCNAC...).
 FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
 FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
 FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
 FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
 FT CARBOHYD 329 329 O-LINKED (POTENTIAL).

 Query Match 8.0%; Score 75.5; DB 1; Length 615;
 Best Local Similarity 20.7%; Pred. No. 6.9;
 Matches 29; Conservative 13; Mismatches 35; Indels 63; Gaps 6;
 QY 1 CP--LQSHALCTCCFQPM-----PDRAREQDPVPAP 32
 DB 121 CPCHLTGNHCCKEFCFQLLAFHFKNEIWTYRTQAAVARCQCKGPDACQ-----RLAS 175
 QY 33 QCAV-----CLQPFCHLYWGCTRTGCGCLAPFCELNIGDKCLDGLVNNNSYSDI 84
 DB 176 QACRTNPLCHGGRCLEVEGHRLCHCP---VGYTGFCFCDVDTKASCYDG----- 220
 QY 85 LKNYLATRLGLTWKNMLTSL 104
 DB 221 -----RGLSTRGLARTL 233

RESULT 11

FA12 HUMAN STANDARD; PRT; 615 AA.

P00748; P78339;

21-JUL-1986 (Rel. 01, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor) (HAF).

DE GN P12.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

OX [1]

SEQUENCE FROM N.A.

RP MEDLINE=88007593; PubMed=2888762;

RP Cool D.E., McGillivray R.T.A.;

RP "Characterization of the human blood coagulation factor XII gene. Intron/exon gene organization and analysis of the 5'-flanking region.";

RT J. Biol. Chem. 262:13662-13673 (1987).

RL [2]

SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.

RP Rieder M.J., Armel T.Z., Carlington D.P., Ozuna M., Kuidanek S.A., Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;

RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

RL [3]

SEQUENCE OF 4-615 FROM N.A.

RP MEDLINE=86176794; PubMed=3754331;

RP Tripidi M., Citarella F., Guida S., Galeffi P., Fantoni A., Cortese R.;

RA "cDNA sequence coding for human coagulation factor XII (Hageman).";

RL Nucleic Acids Res. 14:3146-3146 (1986).

RL [4]

SEQUENCE OF 14-615 FROM N.A.

RP MEDLINE=86033830; PubMed=3877053;

RP Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D., McGillivray R.T.A.;

RA "Characterization of human blood coagulation factor XII cDNA. Prediction of the primary structure of factor XII and the tertiary structure of beta-factor XIIa.";

RT J. Biol. Chem. 260:13666-13676 (1985).

RL [5]

SEQUENCE OF 146-615 FROM N.A.

RP MEDLINE=86216049; PubMed=3011063;

RA Que B.G., Davie E.W.;

RA "Characterization of a cDNA coding for human factor XII (Hageman factor).";

RL Biochemistry 25:1525-1528 (1986).

RL [6]

SEQUENCE OF 20-379

RP MEDLINE=85182874; PubMed=3886654;

RA McMullen B.A., Fujikawa K.;

RA "Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated Hageman factor).";

RT J. Biol. Chem. 260:5328-5341 (1985).

RL [7]

SEQUENCE OF 354-362 AND 373-615.

RP MEDLINE=83291041; PubMed=6604055;

RA Fujikawa K., McMullen B.A.;

RA "Amino acid sequence of human beta-factor XIIa.";

RT J. Biol. Chem. 258:10924-10933 (1983).

RL [8]

SEQUENCE OF 561-615 FROM N.A.

RP TISSUE=Blood;

RC MEDLINE=96133302; PubMed=8528215;

RA Schloesser M., Hofferbert S., Bartz U., Lutz G., Lammle B., Engel V.;

RA "The novel acceptor splice site mutation 11396(G-->A) in the factor XII gene causes a truncated transcript in cross-reacting material negative patients.";

RT Hum. Mol. Genet. 4:1235-1237 (1995).

RL [9]

RESULT 10
ID FAAA_HUMAN STANDARD; PRT; 419 AA.
AC P16930;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fumarylacetoacetase (EC 3.7.1.2) (Fumarylacetoacetate hydrolase)
DE (Beta-diketone) (FAA).
GN FAH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=91150763; PubMed=1998338;
RA Phaneuf D., Labelle Y., Bernube D., Arden K., Cavenee W., Gagne R.,
RA Tanguay R.M.;
RT "Cloning and expression of the cDNA encoding human
RT fumarylacetoacetate hydrolase, the enzyme deficient in hereditary
RT tyrosinemia: assignment of the gene to chromosome 15";
RL Am. J. Hum. Genet. 48:525-535 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
EX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Iqbalano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.N.,
RA Vallaloon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Green E.D., Dickinson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Myers R.M.,
RA Butterfield Y.S.N., Grimwood J., Schmutz J., Myers R.M.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A., Smallus D.E.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE OF 71-419 FROM N.A.
RC TISSUE=Liver;
EX MEDLINE=90245581; PubMed=2336361;
RA Agteribbe E., van Faassen H., Hartog M.V., Reversma T.,
RA Taanman J.-W., Pannkoek H., Evers R.F., Welling G.M., Berger R.;
RT "Nucleotide sequence of cDNA encoding human fumarylacetoacetase";
RL Nucleic Acids Res. 18:1887-1887 (1990).
RN [4]
RP REVIEW ON VARIANTS.
EX MEDLINE=97255958; PubMed=9101289;
RA St Louis M., Tanguay R.M.;
RT "Mutations in the fumarylacetoacetate hydrolase gene causing
RT hereditary tyrosinemia type I: overview";
RL Hum. Mutat. 9:291-299 (1997).
RN [5]
RP VARIANT TYROSINEMIA TYPE I ILR-16.
EX MEDLINE=93016844; PubMed=1401056;
RA Phaneuf D., Lambert M., Laframboise R., Mitchell G., Lettre F.,
RA Tanguay R.M.;
RT "Type I hereditary tyrosinemia. Evidence for molecular heterogeneity
RT and identification of a causal mutation in a French Canadian
RL J. Clin. Invest. 90:1185-1192 (1992).
RN [6]
RP VARIANT TYROSINEMIA TYPE I ASP-134.
RX MEDLINE=93372857; PubMed=8364576;
RA Labelle Y., Phaneuf D., Leclerc B., Tanguay R.M.;
RT "Characterization of the human fumarylacetoacetate hydrolase gene and
RT identification of a missense mutation abolishing enzymatic
RT activity";
RL Hum. Mol. Genet. 2:941-946 (1993).
RN [7]
RP VARIANT TYROSINEMIA TYPE I GLY-166.
EX MEDLINE=93306316; PubMed=8318997;
RA Grompe M., Al-Dhalimy M.;
RT "Mutations of the fumarylacetoacetate hydrolase gene in four patients
RT with tyrosinemia, type I";
RL Hum. Mutat. 2:85-93 (1993).
RN [8]
RP VARIANT TYROSINEMIA TYPE I VAL-233.
EX MEDLINE=95029364; PubMed=7942842;
RA Rootwelt H., Berger R., Gray G., Kelly D.A., Coskun T.,
RA Kvittingen E.A.;
RT "Novel splice, missense, and nonsense mutations in the
RT fumarylacetoacetase gene causing tyrosinemia type I";
RL Am. J. Hum. Genet. 55:653-658 (1994).
RN [9]
RP VARIANT TYROSINEMIA TYPE I TRP-341.
EX MEDLINE=95067974; PubMed=7977370;
RA Rootwelt H., Brodtkorb E., Kvittingen E.A.;
RT "Identification of a frequent pseudodeficiency mutation in the
RT fumarylacetoacetase gene, with implications for diagnosis of
RT tyrosinemia type I";
RL Am. J. Hum. Genet. 55:1122-1127 (1994).
RN [10]
RP VARIANTS TYROSINEMIA TYPE I ASP-134 AND LEU-342.
EX MEDLINE=94274193; PubMed=8005583;
RA Rootwelt H., Chou J., Gahl W.A., Berger R., Coskun T., Brodtkorb E.,
RA Kvittingen E.A.;
RT "Two missense mutations causing tyrosinemia type I with presence and
RT absence of immunoreactive fumarylacetoacetase";
RL Hum. Genet. 93:615-619 (1994).
RN [11]
RP VARIANTS TYROSINEMIA TYPE I SER-337 AND GLY-381.
EX MEDLINE=95276752; PubMed=757089;
RA St Louis M., Poudrier J., Phaneuf D., Leclerc B., Laframboise R.,
RA Tanguay R.M.;
RT "Two novel mutations involved in hereditary tyrosinemia type I";
RL Hum. Mol. Genet. 4:319-320 (1995).
RN [12]
RP VARIANT TYROSINEMIA TYPE I GLY-234.
EX MEDLINE=96055520; PubMed=7550234;
RA Hahn S.H., Kraenewich D., Brantly M., Kvittingen E.A., Gahl W.A.;
RT "Heterozygosity for an exon 12 splicing mutation and a W234G missense
RT mutation in an American child with chronic tyrosinemia type I";
RL Hum. Mutat. 6:66-73 (1995).
RN [13]
RP VARIANTS TYROSINEMIA TYPE I ARG-193 AND VAL-369.
EX MEDLINE=96140750; PubMed=8557261;
RA Floos van Amstel J.K., Bergman A.J.I.W., van Beurden E.A.C.M.,
RA Roijers J.F.M., Peelen T., van den Berg I.E.T., Poll-The B.T.,
RA Kvittingen E.A., Berger R.;
RT "Hereditary tyrosinemia type I: novel missense, nonsense and splice
RT consensus mutations in the human fumarylacetoacetate hydrolase gene;
RT variability of the genotype-phenotype relationship";
RL Hum. Genet. 97:51-59 (1996).
RN [14]
RP VARIANTS TYROSINEMIA TYPE I ASP-158; LEU-261; HIS-405 AND SER-366 DEL.
EX MEDLINE=98295572; PubMed=9633815;
RA Bergman A.J.I.W., van den Berg I.E.T., Brink W., Poll-The B.T.,
RA Floos van Amstel J.K., Berger R.;
RT "Spectrum of mutations in the fumarylacetoacetate hydrolase gene of
RT tyrosinemia type I patients in northwestern Europe and Mediterranean
RL countries";
RL Hum. Mutat. 12:19-26 (1998).
CC -I- CATALYTIC ACTIVITY: 4-fumarylacetoacetate + H(2)O = acetoacetate +
CC fumarate.
CC -I- PATHWAY: Phenylalanine catabolism; sixth step.

Nucleic Acids Res. 26:477-485(1998).

[4]

CHARACTERIZATION.

MEDLINE=98344072; PubMed=9677405;

Gerik K.J., Li X., Pautz A., Burgers P.M.;

"Characterization of the two small subunits of Saccharomyces cerevisiae DNA polymerase delta.";

J. Biol. Chem. 273:19747-19755(1998).

-1- FUNCTION: REQUIRED FOR REPLICATION OF THE LEADING DNA STRAND AND FOR COMPLETION OF LAGGING STRAND SYNTHESIS.

-1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + (DNA) (N).

-1- SUBUNIT: HETEROTRIMER WITH SUBUNITS OF 125 KDA, 58 KDA AND 55 KDA.

-1- SUBCELLULAR LOCATION: Nuclear (By similarity).

-1- SIMILARITY: Belongs to the DNA polymerase delta/II small subunit family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D50324; BAA08859.1; -

EMBL; X87611; CAA60928.1; -

EMBL; Z49506; CAA89528.1; -

PIR; S55194; S55194.

GeneOnline; 141842; -

SGD; S0003766; HYS2.

transferase; DNA-directed DNA polymerase; DNA replication;

Nuclear protein.

CONFLICT 156 156 L -> H (IN REF. 1).

CONFLICT 465 465 S -> N (IN REF. 1).

SEQUENCE 487 AA; 55296 MW; F9E200BF97A2C07 CRC64;

Query Match 8.1%; Score 76.5; DB 1; Length 487;

Best Local Similarity 27.3%; Pred. No. 4.2;

Matches 30; Conservative 19; Mismatches 52; Indels 9; Gaps 4;

54 CYGCLAPCEINLGDKLDGLVNNNSYESDILKNYLATRGLTWKNML-TESLVALQSGVP 112

104 CVGSI--YCEMKYKFNVLVDINDTYGAPDTKSYTDKEGSDLEIMLEDESGRVLVGGP 161

113 LLSDVRYVGTDLVLCYCCGLR----SFRLEITYQRNIPASELPVAVTSRP 158

162 IRSPTPIG--VVVGLGWEAEAGTFQVLDICYPTPLPQNPPAPIATCP 209

RESULT 9

KDGE HUMAN STANDARD; PRT; 567 AA.

AC P52429; O9UK03;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Diacylglycerol kinase, epsilon (EC 2.7.1.107) (Diacylceride kinase)

DE (DGK-epsilon) (DAG kinase epsilon).

GN DGKE OR DAGKS

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

TISSUE=umbilical vein endothelial cells;

MEDLINE=96215320; PubMed=8626589;

Tang W., Bunting M., Zimmerman G.A., McIntyre T.M., Prescott S.M.;

"Molecular cloning of a novel human diacylglycerol kinase highly selective for arachidonate-containing substrates.";

J. Biol. Chem. 271:10237-10241(1996).

[2]

SEQUENCE OF 1-154 FROM N.A.

MEDLINE=20035825; PubMed=10571048;

Tang W., Bardin S., Bhattacharya S.S., Prescott S.M.;

"Characterization of the human diacylglycerol kinase epsilon gene and its assessment as a candidate for inherited retinitis pigmentosa.";

Gene 239:185-192(1999).

-1- FUNCTION: Highly selective for arachidonate-containing species of diacylglycerol (DAG). May terminate signals transmitted through arachidonoyl-DAG or may contribute to the synthesis of phospholipids with defined fatty acid composition.

-1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-diacylglycerol 3-phosphate.

-1- TISSUE SPECIFICITY: Expressed predominantly in testis.

-1- SIMILARITY: Belongs to the eukaryotic diacylglycerol kinase family.

-1- SIMILARITY: Contains 2 zinc-dependent phorbol-ester and DAG binding domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U49379; AAC50497.1; -

EMBL; AF136745; AAD45666.1; -

Gene; HGNC:2852; DGKE.

MTM; 601440; -

GO; GO:0005524; F-ATP binding; TAS.

GO; GO:0004143; F-diacylglycerol kinase activity; TAS.

GO; GO:0008654; P:phospholipid biosynthesis; TAS.

InterPro; IPR002219; DAG PE-bind.

InterPro; IPR000756; DAGKa.

InterPro; IPR001206; DAGKc.

Pfam; PF00130; DAG PE-bind; 2.

Pfam; PF00609; DAGKa; 1.

Pfam; PF00781; DAGKc; 1.

ProDom; PD002939; DAGKa; 1.

ProDom; PD005043; DAGKc; 1.

SMART; SM00109; C1; 2.

SMART; SM00045; DAGKa; 1.

SMART; SM00046; DAGKc; 1.

PROSITE; PS00479; DAG PE BIND DOM 1; 2.

PROSITE; PS50081; DAG PE BIND DOM 2; 2.

Transferrase; Kinase; Phorbol-ester binding; Multigene family;

Transmembrane; Repeat.

TRANSMEM 22 42 POTENTIAL.

TRANSMEM 435 456 POTENTIAL.

FT DOMAIN 60 108 PHORBOL-ESTER AND DAG BINDING 1.

FT DOMAIN 125 177 PHORBOL-ESTER AND DAG BINDING 2.

FT DOMAIN 217 350 CATALYTIC-A (POTENTIAL).

FT DOMAIN 369 524 CATALYTIC-B (POTENTIAL).

SQ SEQUENCE 567 AA; 63927 MW; BC33AD15FB4D0B4 CRC64;

Query Match 8.1%; Score 76; DB 1; Length 567;

Best Local Similarity 21.6%; Pred. No. 5.6;

Matches 36; Conservative 21; Mismatches 54; Indels 56; Gaps 10;

3 LOGSHALCTCCFQMPD---RRAER-----EODPVAPOQ-----CAVCL 39

82 LQG--AFDCGGLRVDEGCLRKADKRFQCKEIMLKNDTKVLDAMPHHWIRGNVPLCSYM 139

40 QPCHLYWGCTRTGC-YGCLAPCEINLGDKLDGLVNNNSYESDILKN-----YLAT- 91

140 --VCKQCCGQPKLCDYRCI--WCQKTVHDECKNLSKNEKCDGFGFKNLIIPPSYLTSI 195

92 -----RGLTWKNN--LTESLVALQSGVPLLSDYRV 119

196 NQMKDKKTDYEVLASLKGKQWTPLIILANSRSGTNNNGELLGEFRI 242

```
QY 129 CG 130
Db 275 CG 276

RESULT 7
ID T1B MOUSE STANDARD; PRT; 834 AA.
AC Q62318; P70391;
DT 15-JUL-1998 (Rel. 36, Created)
DE 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription intermediary factor 1-beta (TIF1-beta) (Tripartite motif
protein 28) (KRAB-A interacting protein) (KRIP-1).
GN TRIM28 OR TIF1B OR KRIP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=9713329; PubMed=8978696;
RA le Douarin B., Nielsen A.L., Garnier J.M., Ichinose H., Jeanmougin F.,
RA Lesson R., Chambon P.;
RT "A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic
RT control of transcription by nuclear receptors.";
RL EMBO J. 15:6701-6715(1996).
RN [2]
SEQUENCE FROM N.A.
RP TTSUB=Kidney;
RX MEDLINE=97140325; PubMed=8986806;
RA Kim S.-S., Chen Y.-M., O'Leary E., Witzgall R., Vidal M.,
RA Bonventre J.V.;
RT "A novel member of the RING finger family, KRIP-1, associates with
RT the KRAB-A transcriptional repressor domain of zinc finger
RT proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:15299-15304(1996).
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=20400347; PubMed=10940561;
RA Cammas F., Garnier J.-M., Chambon P., Losson R.;
RT "Correlation of the exon/intron organization to the conserved domains
RT of the mouse transcriptional corepressor TIF1beta.";
RL Gene 253:231-235(2000).
CC -1- FUNCTION: Forms a complex with a KRAB-domain transcription factor
CC and increases the efficiency of KRAB-mediated repression by
CC recruiting SETDB1 to histone H3 (By similarity).
CC -1- SUBUNIT: Interacts with SETDB1 and CBX3 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: Contains 2 B box-type zinc fingers.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X99644; CAA67963.1; -
CC EMBL; U67303; AAB1722.1; -
CC EMBL; AF230878; AAG02638.1; -
CC MGJ; MGJ109274; Trim28.
CC GO; GO:0000785; C:chromatin; IDA.
CC InterPro; IPR003649; Bbox_C.
CC InterPro; IPR001487; BboxDomain.
CC InterPro; IPR000315; Znf_Bbox.
CC InterPro; IPR001965; Znf_PHD.
CC InterPro; IPR001841; Znf_Ring.
CC Pfam; PF00628; PHD; 1.

Query Match 8.3%; Score 78; DB 1; Length 834;
Best Local Similarity 19.9%; Pred. No. 5.4;
Matches 40; Conservative 18; Mismatches 59; Indels 84; Gaps 8;

QY 10 CTCFQPMEDRAERBQDPVPAQCAVCLQPFCHLYWGCTGTCYGCCLAPFCELNLDG- 68
Db 66 CGVCRE-----RLRPEDRPLLP-----CLHSAC-----SACLGPFAPAAANNNGDG 107
QY 69 -KCLDG-----VLNNNSYESDILKNYL-----ATRG 93
Db 108 GSAGGAMVDCPVCKQCYKDIENFMRDGSKASSDSQDANQCCTSCEDNAPATSYC 167
QY 94 LTKNMLTFSVALQRGVFLLSDY--RVTG-----DT 123
Db 168 VECSEPLCETCEAQRVKYTKDHTVSTGTPAKTRDGERIVYCNVHKHEPLVLFCESCDT 227

RESULT 8
ID DPD2 YEAST STANDARD; PRT; 487 AA.
AC P46957;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase delta small subunit (EC 2.7.7.7).
GN POL31 OR HYS2 OR HUS2 OR SDP5 OR YJR006W OR J1427 OR YJR83.7.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
RA Sugimoto K., Sakamoto Y., Matsumoto K.;
RN Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RL [2]
SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RA de Haan M., Smits P.H.M., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=98083176; PubMed=9421503;
RA Hashimoto K., Nakashima N., Ohara T., Maki S., Sugino A.;
RT "The second subunit of DNA polymerase III (delta) is encoded by the
RT HYS2 gene in Saccharomyces cerevisiae.";
```

DR EMBL; AEO16842; AA070124.1; -.
DR StyGene; SG77777; lpcA.
DR HAMAP; MF 00067; -; 1.
DR InterPro; IPR004515; GnhA.
DR InterPro; IPR001347; SIS.
DR Pfam; PF01380; SIS; 1.
DR TIGRFAHs; TIGR00441; gnhA; 1.
KW isomerase; lipopolysaccharide biosynthesis; Complete proteome.
SQ SEQUENCE 192 AA; 20896 MW; 0ABFCBCECD6786A4 CRC64;
Query Match 8.5%; Score 80; DB 1; Length 192;
Best Local Similarity 30.6%; Pred. No. 0.66;
Matches 22; Conservative 15; Mismatches 27; Indels 8; Gaps 3;
QY 83 DILKXNYLATGLTWKMLTE--SLVALQGVFLISD-YRVTGDTVLC-----YCCGLRSF 134
DB 4 LLIRNELNEAETLANFLKODANIHQRAAVLLADSFKAGGKVLSCNGSGSHCDAMHFA 63
QY 135 RELTYQYRONIP 146
DB 64 EELTGVRNRP 75
RESULT 6
SLI3 HUMAN STANDARD; PRT; 279 AA.
AC Q14192; Q13229; Q13644; Q9P294;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Skeletal muscle LIM-protein 3 (SLIM3) (LIM-domain protein DRAL)
DE (Four and a half LIM domains protein 2) (FHL-2).
GN FHL2 OR SLIM3 OR DRAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RC TIGSUS=Skeletal muscle;
RX MEDLINE=97294674; PubMed=9150430;
RA Genini M., Schwalbe P., Scholl F.A., Remppis A., Mattei M.-G.,
RA Schaefer B.W.;
RT "Subtractive cloning and characterization of DRAL, a novel LIM-domain
RT protein down-regulated in rhabdomyosarcoma."
RL DNA Cell Biol. 16:433-442(1997).
RN [2]
SEQUENCE FROM N.A.
RC TIGSUS=Heart;
RX MEDLINE=98248917; PubMed=9573400;
RA Chan K.K., Teui S.K.W., Lee S.M.Y., Luk S.C.W., Liew C.C., Fung K.P.,
RA Waye M.M.Y., Lee C.Y.;
RT "Molecular cloning and characterization of FHL2, a novel LIM domain
RT protein preferentially expressed in human heart."
RL Gene 210:345-350(1998).
RN [3]
SEQUENCE FROM N.A.
RC TIGSUS=Leukocyte;
RX MEDLINE=20458893; PubMed=11001931;
RA Tanahashi H., Tabira T.;
RT "Alzheimer's disease-associated presenilin 2 interacts with DRAL, an
RT LIM-domain protein."
RL Hum. Mol. Genet. 9:2281-2289(2000).
RN [4]
SEQUENCE FROM N.A.
RC TIGSUS=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
SEQUENCE OF 127-279 FROM N.A.
RN TISSUE=Heart muscle;
RC MEDLINE=96354835; PubMed=8753811;
RX Morgan M.J., Madgwick A.J.A.;
RT "Slim defines a novel family of LIM-proteins expressed in skeletal
RT muscle."
RL Biochem. Biophys. Res. Commun. 225:632-638(1996).
CC -!- TISSUE SPECIFICITY: Expressed only in skeletal muscle.
CC -!- SIMILARITY: Contains 4 LIM zinc-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L42176; AAC85333.1; -.
DR EMBL; U29332; AAC52073.1; -.
DR EMBL; AB038794; BAA92253.1; JOINED.
DR EMBL; AB038991; BAA92253.1; JOINED.
DR EMBL; AB038992; BAA92253.1; JOINED.
DR EMBL; AB038793; BAA92253.1; JOINED.
DR EMBL; AB038792; BAA92253.1; JOINED.
DR EMBL; BC014397; AAH14397.1; -.
DR EMBL; U60117; AAC50794.1; -.
DR PIR; JC6565; JC6565.
DR Genew; HGNC:3703; FHL2.
DR MIM; 602633; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR InterPro; IPR001781; LIM.
DR Pfam; PF0412; LIM; 4.
DR ProDom; PD000094; LIM; 4.
DR SMART; SM00132; LIM; 4.
DR PROSITE; PS00478; LIM DOMAIN 1; 4.
DR PROSITE; PS00023; LIM DOMAIN 2; 4.
KW Repeat; LIM domain; Metal-binding; Zinc; Zinc-finger.
FT ZN FING 7 31 GATA-LIKE (POTENTIAL).
FT DOMAIN 40 92 LIM 1.
FT DOMAIN 101 153 LIM 2.
FT DOMAIN 162 212 LIM 3.
FT DOMAIN 221 275 LIM 4.
FT CONFLICT 167 167 M -> G (IN REF. 1).
FT CONFLICT 167 167 M -> K (IN REF. 3 AND 4).
SQ SEQUENCE 279 AA; 32196 MW; DBDC77C9BD6C4BBA CRC64;
Query Match 8.4%; Score 79; DB 1; Length 279;
Best Local Similarity 23.8%; Pred. No. 1.3;
Matches 29; Conservative 10; Mismatches 51; Indels 32; Gaps 5;
QY 9 LTCFCFQMPDRAREQDPFVAPQCAVCLQPFCHLTGCTTCYCGCLAPFCEINLGD 68
DB 187 VCTACRKQLSGQRTARD-FAVCLNCFCDLY-----AKKCACTNFIISGLG-GT 234
QY 69 KCLDGLVNNNSVESDILKNYLATGLTWKMLTESLVALQGVFLISDVRVTGDTVLCYC 128
DB 235 KVIS--FERQHNDCFNCKKCSLSIVGRGFLTER-----DDLCPD 274


```

RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Pancreas, and Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Aeshburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Type III membrane protein (Potential).
CC -!- INDUCTION: By interferons.
CC -!- SIMILARITY: Belongs to the TMEM7 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AJ251364; CAC13976.1; --
DR EMBL; AK007477; BAB35057.1; --
DR EMBL; AK018021; BAB31041.1; --
DR MGD; MGI:1915025; S830458KL6R1x.
KW Transmembrane.
FT TRANSMEM 228 248 POTENTIAL.
FT CONFLICT 185 185 K -> N (IN REF. 2; BAB31041).
FT CONFLICT 209 209 F -> L (IN REF. 2; BAB31041).
FT SEQUENCE 249 AA; 283391 MW; 2C79B36ED6F042D4 CRC64;
SQ
Query Match 8.6%; Score 81; DB 1; Length 249;
Best Local Similarity 22.2%; Pred. NO. 0.7;
Matches 37; Conservative 21; Mismatches 71; Indels 38; Gaps 6;
QY 17 MPDREA---EREQDPVAPQCAVCLQPF-----CHLY-----WGCTR 51
DB 36 VPDGALGWRHQQTIVLGRFQSCRCRSWTSAQWMLCHMYPDTLKSQQAWEIFGQKC 95
QY 52 TGCYGLAPFCELNLGDKLGVLNNSYSDILKNYLATRLGTLWKNMLTSLVALQGV 111
DB 96 QKCFGC-QETFPKSTBEIKRILNN-----LVNVIQRYGHRKIALTSNLSGKV 146
QY 112 FLLSDYRVGTGVTLVLCYCGLSFRLTQYQRNPASELPVAVTSRP 158
DB 147 TLDGPH-----DTRNCEACLSHGRCALAHKVPKPPSPSPKXSSP 189
RESULT 4
LPCA_ECOLI
ID LPCA_ECOLI STANDARD; PRT; 192 AA.
AC F51001;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoheptose isomerase (EC 5.-.-.-).
GN LPCA OR GWEA OR TFRA OR B0222 OR C0372 OR Z0280 OR ECS0249 OR SF0272
GN OR S0293
OS Escherichia coli,

```

```

OS Escherichia coli O6,
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria, Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562, 217992, 83334, 623;
[1]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / W3110;
RC MEDLINE=96216460; PubMed=8631969;
RA Brooke J.S., Valvano M.A.;
RT "Biosynthesis of inner core lipopolysaccharide in enteric bacteria
RT identification and characterization of a conserved phosphoheptose
RT isomerase.";
RL J. Biol. Chem. 271:3608-3614(1996).
[2]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RC Olmori H.; STRAIN=K12;
RA Olmori H.; (OCT-1995) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MGI655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[4]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / W3110;
RA Yamamoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Takemoto Y., Inokuchi H., Miki T., Hataeda E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RC Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Pederspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
[6]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RC MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
[7]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[8]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;

```

```
CC -----
DR ENBL; AF025441; AAC39561.1; ALT_INIT.
DR ENBL; BC015050; AAH15050.1; -.
DR MIM; 606020; -.
DR GO; GO:0005515; P:protein binding; TAS.
DR GO; GO:0007154; P:cell communication; NAS.
SQ SEQUENCE 229 AA; 24691 MW; 0EBD4006193A3106 CRC64;

Query Match          9.0%; Score 85; DB 1; Length 229;
Best Local Similarity 26.6%; Pred. No. 0.25;
Matches 38; Conservative 15; Mismatches 52; Indels 38; Gaps 6;

QY 26 QDPRVAPQCAVC---LQPPCHLYWGCTRTGCGYCLAPFCELNIGDKCLDGLVNNNSYES 82
DB 71 QPERCAVFQCAQCHAVLADSLVHLADLSR-----SLGAVFVSRTNNVLEA 117
QY 83 DILKNYLAT-RLGTWKML-----TESLVALQGVFLISDYRVGTGTVLC 126
DB 118 PFLVIGIEGLKSGTY-NLLFCGSGIPGVGHLYSTHAAALAGHFCLS-----SDKMYC 171
QY 127 YCCGLRSFRLTYQYRONIPASE 149
DB 172 YLLKTKAIVNASEMDIQNVPLSE 194

RESULT 2
Z451 MOUSE
ID_Z451_MOUSE STANDARD; PRT; 1056 AA.
AC QSCQF7; 2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Zinc finger protein 451.
GN ZNF451.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaudo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Vitaranto R., Wagner L., Warstead C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki K., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
CC -!- FUNCTION: May be involved in transcriptional regulation.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
```

```
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 12 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AK030088; BAC26778.1; -.
DR MGD; MGI:2138298; AI596398.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 7.
DR SMART; SM00355; Znf_C2H2; 12.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 5.
DR Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT ZN_FING 169 195 C2H2-TYPE 1.
FT ZN_FING 212 234 C2H2-TYPE 2 (DEGENERATE).
FT ZN_FING 253 277 C2H2-TYPE 3.
FT ZN_FING 315 338 C2H2-TYPE 4 (ATYPICAL).
FT ZN_FING 362 385 C2H2-TYPE 5.
FT ZN_FING 494 517 C2H2-TYPE 6.
FT ZN_FING 527 550 C2H2-TYPE 7.
FT ZN_FING 604 629 C2H2-TYPE 8 (ATYPICAL).
FT ZN_FING 634 657 C2H2-TYPE 9.
FT ZN_FING 665 688 C2H2-TYPE 10.
FT ZN_FING 751 774 C2H2-TYPE 11.
FT ZN_FING 787 810 C2H2-TYPE 12.
SQ SEQUENCE 1056 AA; 120069 MW; 7D8FBC0B50ECA622 CRC64;

Query Match          8.6%; Score 81.5; DB 1; Length 1056;
Best Local Similarity 25.0%; Pred. No. 3.1;
Matches 38; Conservative 17; Mismatches 52; Indels 45; Gaps 8;

QY 35 CAVCLQPPFC-----HLVWGCTRTG-----CYCLAPPFCELNIGDKCLD 72
DB 214 CAVCYEHFVTCQYKNDHL---LSRTAAADGHSNLSLPQIQCYACPOCFLLFSTKDECLK 270
QY 73 GVLNNNSYESD--ILKNYLATRTGNTKMLTESLVALQGVFLISDYRVGTGTVLCYCC- 129
DB 271 HMTKKNHFGQSKLSDNKGRTARPISPFSFAKRLVSLCKDVPF-----QVKCVACH 321
QY 130 -GLRSFRLTYQYR---QNIPASBELPVAVTSR 157
DB 322 QTLRSHMELTAHFRVRCQVAG---PVALAEK 349

RESULT 3
ID IF28_MOUSE
AC Q9ER80; Q9D3D6; STANDARD; PRT; 249 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 28 kDa interferon responsive protein.
GN IFRG28.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2;
RT Meritet J.F., Dron M., Tovey M.;
RT "Characterization of ifrg28 and ifrg28, two newly identified
RT interferon responsive gene."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:37:32 ; Search time 5.60929 Seconds
(without alignments)
1540.951 Million cell updates/sec

Title: US-10-046-046-2_COPY_476_641

Perfect score: 943
Sequence: 1 CPGQSHALCTCCFCQMPDR.....ASLPVAVTSRDPYWGNC 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	9.0	229	1 OIP5 HUMAN	O43482 homo sapien
2	81.5	8.6	1056	1 Q8COP7	mus musculus
3	81	8.6	249	1 IF28 MOUSE	Q9er80 mus musculus
4	80	8.5	192	1 LPCA ECOLI	P51001 escherichia
5	80	8.5	192	1 LPCA SALTY	Q8Xg28 salmonella
6	79	8.4	279	1 SLI3 HUMAN	Q14192 homo sapien
7	78	8.3	834	1 TP1B MOUSE	O62318 mus musculus
8	76.5	8.1	487	1 DP22 YEAST	P46957 saccharomyc
9	76	8.1	567	1 KGE2 HUMAN	P52429 homo sapien
10	75.5	8.0	419	1 FAAA HUMAN	P16930 homo sapien
11	75.5	8.0	615	1 FA12 HUMAN	P00748 homo sapien
12	75	8.0	193	1 LPCA YERPE	Q82BY7 yersinia pe
13	75	8.0	279	1 SLI3 MOUSE	O70433 mus musculus
14	75	8.0	279	1 SLI3 RAT	O35115 rattus norv
15	75	8.0	465	1 OPCA NOSPU	P48971 nostoc punc
16	75	8.0	1180	1 RNT1 DROME	Q9VY83 drosophila
17	75	8.0	1458	1 PHIX RABIT	Q05017 oryctolagus
18	74.5	7.9	603	1 FA12 CAVPO	Q04962 cavia porce
19	73.5	7.8	419	1 FAAA MOUSE	P35505 mus musculus
20	73.5	7.8	419	1 FAAA RAT	P25093 rattus norv
21	73	7.7	603	1 PGH2 CHICK	P27607 gallus gall
22	72.5	7.7	835	1 TP1B HUMAN	Q13263 homo sapien
23	72	7.6	193	1 PHOL1 PHOLI	Q7H7F7 photornadu
24	72	7.6	3075	1 LNA1 HUMAN	P23391 homo sapien
25	71.5	7.6	854	1 NAL6 RAT	Q63035 rattus norv
26	71	7.5	450	1 NRI1 HUMAN	Q9BU40 homo sapien
27	71	7.5	485	1 RNI4 MOUSE	Q91190 mus musculus
28	71	7.5	3718	1 LW25 MOUSE	O61001 mus musculus
29	70.5	7.5	192	1 YD24 AQUAE	O67344 aquifex seo
30	70.5	7.5	351	1 Y876 METJA	Q58286 methanococc
31	70.5	7.5	387	1 SGCA MOUSE	P82350 mus musculus
32	70.5	7.5	1135	1 VGLM TSWV1	P36291 tomato spot
33	70	7.4	1093	1 RNT1_NEUCR	Q9heh1 neurospora

34	69.5	7.4	63	1 KURT PARTR	P58910 parabuthus
35	69.5	7.4	2139	1 CRB DROME	P10040 drosophila
36	69	7.3	305	1 FX34 HUMAN	Q9AWN3 homo sapien
37	68.5	7.3	245	1 YPC2 ECOLI	P18128 escherichia
38	68.5	7.3	456	1 NRI1 CHICK	Q90ZD5 gallus gall
39	68.5	7.3	491	1 2ASR MOUSE	Q9Z176 mus musculus
40	68.5	7.3	780	1 OSTA YERPE	O8ZIK3 yersinia pe
41	68.5	7.3	1122	1 TEST MOUSE	O70372 mus musculus
42	68.5	7.3	1341	1 YI78 YEAST	Q05854 saccharomyc
43	68.5	7.3	2715	1 GL56_PAPPR	P13837 paramascium
44	68	7.2	342	1 ARGC CANJE	Q9PIS0 campylobact
45	68	7.2	550	1 SYN_CHLTR	O84035 chlamydia t

ALIGNMENTS

RESULT 1
OIP5_HUMAN STANDARD; PRT; 229 AA.
AC O43482: Q96BX7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB Opa-interacting protein 5.
GN OIP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP MEDLINE=98125741; PubMed=9466265;
RX Williams J.M., Chen G.-C., Zhu L., Rest R.F.;
RT "Using the yeast two-hybrid system to identify human epithelial cell
RT proteins that bind gonococcal Opa proteins: intracellular gonococci
RT bind pyruvate kinase via their Opa proteins and require host pyruvate
RT for growth.";
RL Mol. Microbiol. 27:171-186(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udgin T.B., Tothiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gurnatne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield A.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBUNIT: Binds outer membrane protein Opa from Neisseria
CC gonorrhoeae.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

A:Molecule type: DNA
A:Residues: 1-1101 <GEI>
A:Cross-references: EMBL:U39644; NID:g1049339; PID:g1049343; PIDN:AAA80360.1; CESP:T10E1
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:T10E10.4
A:Introns: 93/2; 152/2; 191/3; 209/2; 283/3; 303/1; 399/3; 421/1; 440/1; 465/1; 547/3; 7
Query Match 8.0%; Score 75.5; DB 2; Length 1101;
Best Local Similarity 20.3%; Pred. No. 54;
Matches 35; Conservative 27; Mismatches 55; Indels 55; Gaps 9;
QY 1 CPLQGSHALCTCCFQMPDRAERQDPRAVQQAQVCLQPFCHLYWGCTRTGCGYGLAP 60
Db 558 CGVGSYKTTCCQEI---KLPVCSNGMISQKRCIVAAE--CGLNLECSNGGC--CPIP 610
QY 61 FC-----ELN---LGDKCLDGVLNNSYESDILKNVLTATRGLTAKNMLTESLVA 106
Db 611 FCFNGVTARGRCQVNGCPMGQACMEGLC-----CPLPK---CSNGITSLGICITRLDC 661
QY 107 LQEGV-----FLLSDYRVVTGDTVLCY-----CCGLRS 133
Db 662 GRIGVDCNNGACCPLETCPNNIASSQRCGGCTNCCPVGQTCMNGGCCDLPS 713

Search completed: May 7, 2004, 14:50:10
Job time : 11.5404 secs

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: H04724
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: H04724
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-543 <STO>
A/Cross-references: GB:AB002093; NID:g4887759; PIDN:AA032295.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g31770
A/Map position: 2

Query Match 8.0%; Score 75.5; DB 2; Length 543;
Best Local Similarity 25.3%; Pred. No. 27;
Matches 24; Conservative 10; Mismatches 32; Indels 29; Gaps 5;

QY 10 CTCGQPMPPRAAREQDPRAVQCAVCLQPFCHLYW-----GCTRTGCGYC 57
128 CGICFESY-----TREIARVS-----CGHPYCKTCWAGYITTKIEDGPGCLRV---KC 173

QY 58 LAPCELNGLGDKLDGVLN---NNSYESDILKNYL 89
DB 174 PEPSCSAAVGMDIEDVTETKVNKISYILRSYV 208

RESULT 14
KFHUI2
coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human
N/Alternate names: Hageman factor (activated)
C/Species: Homo sapiens (man)
C/Date: 27-Nov-1985 #sequence_revision 30-Jun-1991 #text_change 08-Dec-2000
C/Accession: A29411; A26814; A00930; A25191; A22248; A21037

R/Cool, D.E.; MacGillivray, R.T.A.
J. Biol. Chem. 262, 13662-13673, 1987
A/Title: Characterization of the human blood coagulation factor XII gene. Intron/exon ge
A/Reference number: A29411; MUID:86007593; PMID:2888762

A/Accession: A29411
A/Molecule type: DNA
A/Residues: 1-615 <COO>
A/Cross-references: GB:M17466; GB:J02807; NID:g180355; PIDN:AA059490.1; PID:g180357

R/Tripodi, M.; Citarella, P.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.
Nucleic Acids Res. 14, 3146, 1986
A/Title: cDNA sequence coding for human coagulation factor XII (Hageman).
A/Reference number: A26814; MUID:86176794; PMID:3754331

A/Accession: A26814
A/Molecule type: mRNA
A/Residues: 4-615 <TRI>
A/Cross-references: GB:M31315; NID:g182291; PIDN:AAA70225.1; PID:g182292

R/Cool, D.E.; Edgell, C.J.; Louis, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R.T.
J. Biol. Chem. 260, 13666-13676, 1985
A/Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the
A/Reference number: A00930; MUID:86033830; PMID:3877053

A/Accession: A00930
A/Molecule type: mRNA
A/Residues: 14-332, 'S', 334-615 <CO2>
A/Cross-references: GB:M11723; NID:g180358; PIDN:AAA51986.1; PID:g180359

R/Que, B.G.; Davie, E.W.
Biochemistry 25, 1525-1528, 1986
A/Title: Characterization of a cDNA coding for human factor XII (Hageman factor).
A/Reference number: A25191; MUID:86216049; PMID:3011063

A/Accession: A25191
A/Molecule type: mRNA
A/Residues: 146-378, 'G', 380-615 <QUE>
A/Cross-references: GB:M3147; NID:g180360; PIDN:AAA70224.1; PID:g180361

R/McMullen, B.A.; Fujikawa, K.
J. Biol. Chem. 260, 5328-5341, 1985
A/Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated Ha
A/Reference number: A22248; MUID:85182674; PMID:3886654

A/Accession: A22248
A/Molecule type: protein
A/Residues: 20-379 <MCM>
R/Fujikawa, K.; McMullen, B.A.
J. Biol. Chem. 258, 10924-10933, 1983
A/Title: Amino acid sequence of human beta-factor XIIa.
A/Reference number: A21037; MUID:83291041; PMID:6604055
A/Accession: A21037
A/Molecule type: protein
A/Residues: 354-362; 373-615 <FUD>
R/Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A/Title: O-linked fucose is present in the first epidermal growth factor domain of factor
A/Reference number: A44606; MUID:92184750; PMID:1544894
A/Contents: annotation; carbohydrate binding site
C/Genetics:
A/Gene: GDB:F12
A/Cross-references: GDB:119892; OMIM:234000
A/Map position: 5q34-5qter
A/Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 463/1; 511/1
C/Complex: factor XII, prekallikrein, and HMW kininogen form a complex bound to anionic s
C/Function:
A/Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasma p
ikrein
A/Pathway: blood coagulation; fibrinolysis
C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C/Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringie; plasma; se
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-372,373-615/Product: coagulation factor XIIa, alpha form #status experimental <A12>
F/47-88/Domain: fibronectin type II repeat homology <F82>
F/98-130/Domain: EGF homology <EG1>
F/135-170/Domain: fibronectin type I repeat homology <1F1>
F/178-209/Domain: EGF homology <EG2>
F/217-295/Domain: kringie homology <KRG>
F/238-356/Region: proline-rich
F/334-362,373-615/Product: coagulation factor XIIa, beta form #status experimental <B12>
F/373-609/Domain: tryptsin homology <TRY>
F/98-110,104-119,121-130,135-163,161-170,178-189,183-198,200-209,217-295,238-277,266-290,
F/109/Binding site: carbohydrate (Thr) (covalent) #status experimental
F/249,433/Binding site: carbohydrate (Asn) (covalent) #status experimental
F/259,305,328,329,337/Binding site: carbohydrate (Thr) (covalent) #status predicted
F/308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F/412,461,563/Active site: His, Asp, Ser #status predicted

Query Match 8.0%; Score 75.5; DB 1; Length 615;
Best Local Similarity 20.7%; Pred. No. 30;
Matches 29; Conservative 13; Mismatches 35; Indels 63; Gaps 6;

QY 1 CP--LOGSHALCTCCFQPM-----PDRAREQDDPRVAP 32
121 CPOHLTGNHCQKCEPEQLLFFHKNWIWYRTEQAAVRCCKGPDACQ-----RLAS 175

QY 33 QQCAV-----CLOPFCHLYWGCTRTGCGCLAPFCFLNLGDKCLDGLVANNNSYESDI 84
176 QACRTNPCLHGGRCLEVEGHRLLCHCP----VGTGFCVDVTKASCYDG-----220

QY 85 LKMYLATRGLTWKNMLTESL 104
221 -----RGLSVYRGLARTTL 233

RESULT 15
TI6840
hypothetical protein TI0E10.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: TI6840
R/Geisel, C.
submitted to the EMBL Data Library, October 1995

A/Description: The sequence of C. elegans cosmid TI0E10.
A/Reference number: Z18588
A/Accession: TI6840
A/Status: preliminary; translated from GB/EMBL/DD5J

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AD0542

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-192 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD08780.1; PID:g16501596; GSPDB:GN00176

C;Genetics:

A;Gene: STY0355

C;Superfamily: phosphoheptose isomerase

Query Match 8.5%; Score 80; DB 2; Length 192;

Best Local Similarity 30.6%; Pred. No. 3.6;

Matches 22; Conservative 15; Mismatches 27; Indels 8; Gaps 3;

QY 83 DILKNYLATRGLTWKNNLTET--SLVALQRGVFLSD-YRVTTGDTVLC-----YCCGLRSF 134

DB 4 DLRNELNEAETLANFLKDDANIHAIQRAAVILLASFRAGGKVLSCGSGSHCDAMHFA 63

QY 135 RELTYQYRONIP 146

DB 64 EELTGYRENRP 75

RESULT 9

T00480

probable RING zinc finger protein A2g34990 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F1913.22

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 12-Feb-1999 #sequence_revision 12-feb-1999 #text_change 23-Mar-2001

C;Accession: T00480; C84763

R;Roundley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.

A;Submitted to the EMBL data Library, April 1998

A;Description: Arabidopsis thaliana chromosome II BAC F1913 genomic sequence.

A;Reference number: Z14160

A;Accession: T00480

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-302 <ROU>

A;Cross-references: EMBL:AC004238; NID:g30333373; PID:g30333394

A;Experimental source: Cultivar Columbia

R;Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umavam, L.; T

euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.;

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana

A;Reference number: AB4420; MUID:20083487; PMID:10617197

A;Accession: C84763

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-302 <STO>

A;Cross-references: GB:AB002093; NID:g30333394; PIDN:AAC12838.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g34990; F1913.22

A;Map position: 2

C;Superfamily: Arabidopsis hypothetical protein F1913.22; RING finger homology

Query Match 8.5%; Score 80; DB 2; Length 302;

Best Local Similarity 22.5%; Pred. No. 5.7;

Matches 45; Conservative 18; Mismatches 53; Indels 84; Gaps 10;

QY 34 QCACVCIQPF-----CHLY-----WGCTRTGCGCLAFCELNLDGKCL--- 71

DB 95 ECACVCEPFEDHETRLMPSCCHVFHADCVSWLSDHSTC-----PLCRVDL---CLQPG 146

QY 72 -----DGVNNNSYESDILKNYLATRGLTWKNNLTET-----LVALQRGVFLSD 116

DB 147 ERSYLNPEPDLVSTNSHLFD-----GVVTRNRPSSRSWSTRLSOCRVSQLISR 197

QY 117 YRVTTGDTVLCYCCGLRSF-----RELTYQ-----YRONIPASELP 151

DB 198 SHSTGSHVQPLNLDRLFTLRLPVEVRQLTKTVDNVAFSQARSSRGVYRSAGSERS 257

QY 152 V-----AVTSRRPCYWGRCN 166

A:Accession: C96516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <STO>
A:Cross-references: GB:AE005173; NID:g5668816; PIDN:ABD46042.1; GSPDB:GN00141
A:Gene: E16N3.15
A:Map position: 1

Query Match 14.6%; Score 137.5; DB 2; Length 473;
Best Local Similarity 25.0%; Pred. No. 4.2e-05;
Matches 45; Conservative 26; Mismatches 88; Indels 21; Gaps 6;

QY 1 CPLQSHALCTCCFQMPDRRAERQDPRAVAPQCAVCLQPFCHLYWGCTRTGCYGCCLAP 60
Db 275 CEHGHAGLQCHLCQGMPPRANLQ-----VFLHCKGCDRPFQGYWS-SENVTGVGSGP 327

QY 61 FCELN-----LQDKLDGV-----LNNNSYSDILKNYLATRLGTWKNMLTESI-----VAL 107
Db 328 VCVRETRFRISERTITRIPIFTHENMRHQDITQRCIAHMEKTVDPVVAEMLRLFNREI 387

QY 108 QGVFLLSYR-VTGTDTLVCYCGLSRPRELTVQRQNPASELPVATSRPDCYWGRCN 166
Db 388 DSRMPLNHAETITASTHVNCDCYDKLVGFLYWFRLTPRNHLPADVAAREDCWGYAC 447

RESULT 3
T22759
hypothetical protein F55H12.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: T22759
R:Doonan, R.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19610
A:Accession: T22759
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-2824 <WIL>
A:Cross-references: EMBL:281091; PIDN:CA03143.1; GSPDB:GN00019; CESP:F55H12.3
A:Experimental source: clone F55H12
C:Genetics:
A:Gene: CESP:F55H12.3
A:Map position: 1
A:Introns: 87/1; 98/1; 126/2; 201/3; 343/3; 406/1; 576/3; 656/1; 825/3; 869/1; 909/1; 967/1; 1755/2; 1800/1; 1850/3; 1896/1; 2003/3; 2035/3; 2082/3; 2119/1; 2144/1; 2200/2; 2277/1; 2279/1
C:Superfamily: LDL receptor ligand-binding repeat homology
F:243-279/Domain: LDL receptor ligand-binding repeat homology <LDL>

Query Match 8.9%; Score 83.5; DB 2; Length 2824;
Best Local Similarity 27.1%; Pred. No. 26;
Matches 55; Conservative 12; Mismatches 51; Indels 85; Gaps 15;

QY 10 CTCCTCFP-----MPDRRAERQDPRAVAPQCAVCLQ-----P 41
Db 1407 CLCNPPCPSTELLQPR--DQIGDCTLNSCQNGVCLQNGMHDCACNEFFIVDRENGTP 1464

QY 42 FCH-----LYWGCTRTGCYGC-----LAPCELN-LQDKC-LDGVLNNSYE-----SD 83
Db 1465 FCPNHLFTNQTRNSGYDCSTGRENALFLCP--NLNFGDFQCYEGQLYNNSYIVLSA 1522

QY 84 ILKNYLATRLGTWKNMLTSLVALQGVFLLSYRVTGTDTLVCYCGLSRPRELTVQRQ 143
Db 1523 GIENAIATN-----LCBSLDTL-----YAVPN-----TFCV-----Q 1549

QY 144 NIPASELPVATSRPDCYWGRCN 166
Db 1550 N-PTSTTP-SIHRCDFCYGGENC 1570

C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C:Accession: T31070
R:Sherwood, D.R.; McClay, D.R.
Development 124; 3393-3574, 1997
A:Title: Identification and localization of a sea urchin Notch homologue: insights into notch homolog - sea urchin (Lytechinus variegatus)
A:Reference number: Z20966; MUID:97454256; PMID:9310331
A:Accession: T31070
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-2531 <SHE>
A:Cross-references: EMBL:AF000634; NID:g2570350; PID:g2570351; PIDN:AA82088.1
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match 8.8%; Score 83; DB 2; Length 2531;
Best Local Similarity 28.6%; Pred. No. 25;
Matches 30; Conservative 9; Mismatches 40; Indels 26; Gaps 6;

QY 5 GSHALCTCCFQMPDRRAERQDPRAVAPQCAVCLQPFCHLYWGCTRTGCYGCCLP 58
Db 840 GYSCRTSGTGGFQGNFCDORNE-----CLFSPCRNGSGCTNLEGSFECSCLP 885

QY 59 ---APFCELNLGDKCLDGVLNNSYSDILKNYLAT--RGLTWKN 98
Db 886 GYDGPICEINI-DECASGPCTNGGICTDLDDYFCSCQRGFTGKN 929

RESULT 5
G64746
phosphoheptose isomerase (EC 5.---) gmha - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C:Accession: G64746
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277; 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G64746
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-192 <BLAT>
A:Cross-references: GB:AE000131; GB:U00096; NID:g1786415; PIDN:AACT3326.1; PID:g1786416;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: gmha
C:Function:
A:Description: synthesis of glyceromannohexose 7-phosphate
A:Pathway: inner core lipopolysaccharide biosynthesis
C:Superfamily: phosphoheptose isomerase
C:Keywords: isomerase

Query Match 8.5%; Score 80; DB 1; Length 192;
Best Local Similarity 30.6%; Pred. No. 3.6;
Matches 22; Conservative 15; Mismatches 27; Indels 8; Gaps 3;

QY 83 DILKNYLATRLGTWKNMLTE--SLVALQGVFLLSD-YRVGTDTVLC-----YCCGLRSF 134
Db 4 DLIRNELNEAETLANFLKDDANIHAQRAAVLLADSFKAGGKVLSCNGSGHCDMHFA 63

QY 135 RELTVQYQENIP 146
Db 64 BELTGRYENRP 75

RESULT 6
A90660
phosphoheptose isomerase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: A90660
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:41:17 ; Search time 9.29039 Seconds
(without alignments)
1718.743 Million cell updates/sec

Title: US-10-048-046-2_COPY_476_641
Perfect score: 943
Sequence: 1 CPLQGSHALCTCCFQMPDR.....ASLPVAVTSRPPDCYWGRC 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: Pir1.*
- 2: Pir2.*
- 3: Pir3.*
- 4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	943	100.0	306	2 T46399	hypothetical prote
2	137.5	14.6	473	2 C96516	F16N3.15 [imported]
3	83.5	8.9	2824	2 T22759	hypothetical prote
4	83	8.8	2531	2 T11070	notch homolog - se
5	80	8.5	192	1 G54746	phosphoheptose iso
6	80	8.5	192	2 A30660	phosphoheptose iso
7	80	8.5	192	2 G85510	phosphoheptose iso
8	80	8.5	192	2 A05420	phosphoheptose iso
9	80	8.5	302	2 T00480	probable RING zinc
10	79	8.4	279	2 J65855	four-and-a-half LI
11	76.5	8.1	487	1 S55194	DNA-directed DNA p
12	75.5	8.0	419	2 A37926	fumarylacetoacetat
13	75.5	8.0	543	2 H84724	probable ARI-like
14	75.5	8.0	615	1 KFHU12	coagulation factor
15	75.5	8.0	1101	2 T16840	hypothetical prote
16	75	8.0	193	2 A80394	phosphoheptose iso
17	75	8.0	1458	2 A56855	adult-specific bru
18	74.5	7.9	419	2 A56825	fumarylacetoacetat
19	74.5	7.9	603	2 S28941	coagulation factor
20	73.5	7.8	362	2 T02605	hypothetical prote
21	73.5	7.8	406	2 T24021	hypothetical prote
22	73.5	7.8	419	2 A40219	fumarylacetoacetat
23	73.5	7.8	419	2 JH0467	fumarylacetoacetat
24	73.5	7.8	679	2 B96599	protein F20N2.12 [
25	73	7.7	603	2 A38630	prostaglandin-endo
26	72	7.6	3075	2 S14458	laminin alpha-1 ch
27	71.5	7.6	483	2 S27880	Nasopressin recept
28	71.5	7.6	4550	2 T18440	hypothetical prote
29	71	7.5	3635	2 T10053	laminin alpha 5 ch

30	70.5	7.5	192	2 E70414	hypothetical prote
31	70.5	7.5	351	2 D64409	iron(III) dicitrat
32	70.5	7.5	387	2 J55556	adhalin - mouse
33	70.5	7.5	421	2 AC1096	hypothetical prote
34	70.5	7.5	1135	1 JQ1928	G2-G1 polypeptide
35	70	7.4	661	2 B96596	hypothetical prote
36	70	7.4	802	2 T24293	hypothetical prote
37	70	7.4	949	2 T24294	hypothetical prote
38	69.5	7.4	320	2 T45909	hypothetical prote
39	69.5	7.4	401	2 T02958	ribulose-bisphosph
40	69.5	7.4	2139	2 A35672	crumbs protein - f
41	69	7.3	302	2 AE0645	probable ROK-famil
42	69	7.3	347	2 T33204	hypothetical prote
43	69	7.3	378	2 T00481	probable RING zinc
44	69	7.3	456	2 T19817	hypothetical prote
45	68.5	7.3	444	2 A55071	hydrogen peroxide-

ALIGNMENTS

RESULT 1

T46399
Hypothetical protein DKFZp434N2420.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46399
R;Ottewaelde, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46399
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-306 <AAA>
A;Cross-references: EMBL:AL137561
A;Experimental source: adult testis; clone DKFZp434N2420
C;Genetics:
A;Note: DKFZp434N2420.1

Query Match	100.0%;	Score 943;	DB 2;	Length 306;
Best Local Similarity	100.0%;	Pred. No. 7e-80;	Mismatches 0;	Indels 0;
Matches 166;	Conservative 0;	Gaps 0;		
Qy	1	CPLQGSHALCTCCFQMPDRRAERQDPRAVAPQCAVCLQPFCHLYWGCTRTGCGYGLAP	60	
Db	118	CPLQGSHALCTCCFQMPDRRAERQDPRAVAPQCAVCLQPFCHLYWGCTRTGCGYGLAP	177	
Qy	61	FCEINLGDKCLDGLVNNNSYESDILKNYLATGLTKWNLTSVALQGVLLSDYRVT	120	
Db	178	FCEINLGDKCLDGLVNNNSYESDILKNYLATGLTKWNLTSVALQGVLLSDYRVT	237	
Qy	121	GDTVLCYCCGLRSFRLTYQYRQNPASLPVAVTSRPPDCYWGRC	166	
Db	238	GDTVLCYCCGLRSFRLTYQYRQNPASLPVAVTSRPPDCYWGRC	283	

RESULT 2

C96516
F16N3.15 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: C96516
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewart, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712

```

; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 495
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-495

Query Match      8.2%; Score 77; DB 14; Length 771;
Best Local Similarity 29.4%; Pred. No. 32;
Matches 20; Conservative 1; Mismatches 29; Indels 18; Gaps 2;

QY 1 CPLLGGSHALCTCCFQPMPPDRRAEREQDPVAPQCCAVCLQPFCHLYWGCTRTG-CYGCILA 59
Db 98 CCCAGTGGCCCC-----ATGACTCTTACCTGATGCTGTGCCAGCCA 140

QY 60 PFCELNLG 67
Db 141 CACAAGAG 148
```

Search completed: May 7, 2004, 15:06:52
Job time : 27.2936 secs

```

RESULT 12
US-10-425-114-62319
; Sequence 62319, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62319
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4666-024-H5_FLI.pep
US-10-425-114-62319

Query Match      8.3%; Score 78.5; DB 12; Length 216;
Best Local Similarity 24.8%; Pred. No. 4.4;
Matches 40; Conservative 11; Mismatches 55; Indels 55; Gaps 10;

QY 2 PLOGSHALCTCCFQMPDRAERQDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAPP 61
DB 74 PTQWSSGICAC-----FDD-----PQSC-----CIGAICP- 98

QY 62 CELNLGDKLDGVLNNNSYSDILKNYLATRGLTWKNMLESVALQRGVFLLSYRVTG 121
DB 99 -----CFLGKNAQFLSGTLAGSCTTHCMLW-GLLT-SLCCVFTGGLVLA---VPG 145

QY 122 DTVLCYCCGLRSFRLTYQYRONIPASELPVA-VTSRPDCY 161
DB 146 SAVACYACGYRSALRTKY-----NLP--EAPCGDLTLHLFCH 180

RESULT 13
US-10-108-260A-3965
; Sequence 3965, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3965
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3965

Query Match      8.2%; Score 77; DB 15; Length 303;
Best Local Similarity 22.3%; Pred. No. 9.9;
Matches 39; Conservative 19; Mismatches 65; Indels 52; Gaps 10;

QY 1 CPLQGS-----ALCTCCFQMPDRAERQDPRVAP----- 32
DB 130 CPVASSDFGQVLECPSCCHKPCSCCKDAWHAEVSCDSQPIVLPTEHREALFGTDAEP 189
QY 33 -QCAVCLQPFCHLYWGCTRTGCGCLAPPCELNGLDKLDGVLNN---NSYSDILKNY 88
DB 190 IKQCPVC-RVYIERNEGCAQWCKNCKHTFCWY-----CLQNLNDIFLRYDKGFCRNK 243

RESULT 14
US-10-142-426-495
; Sequence 495, Application US/10142426
; Publication No. US20040048333A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Godowski, Steven
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C224
; CURRENT APPLICATION NUMBER: US/10/142,426
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 495
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-426-495

Query Match      8.2%; Score 77; DB 12; Length 771;
Best Local Similarity 29.4%; Pred. No. 32;
Matches 20; Conservative 1; Mismatches 29; Indels 18; Gaps 2;

QY 1 CPLQGSALCTCCFQMPDRAERQDPRVAPQCAVCLQPFCHLYWGCTRTG-CYGCCLA 59
DB 98 CCACGTGGCCCC-----ATGACTCTTACTGATGCTGTGCCAGCCA 140

QY 60 PFCEINLG 67
DB 141 CACAAGAG 148

RESULT 15
US-10-123-155-495
; Sequence 495, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
```


US-09-942-052-731

Query Match 9.0%; Score 85; DB 10; Length 229;
Best Local Similarity 26.6%; Pred. No. 0.96;
Matches 38; Conservative 15; Mismatches 52; Indels 38; Gaps 6;

QY 26 QDRVAPQCAVC---LQPFCHLYWGCTRGCGYCLAPFCELNLDGKLDGLVNNNSYES 82
DB 71 QPERCAVFOCAQCHAVLADSVHLADLSR-----SLGAVFSRTNNVLEA 117
QY 83 DILKNYLAT-RGLTWKNML-----TESVALQGVFLLSDYRVGTDTVC 126
DB 118 PFLVGIEGSLKSTY-NLLFCGSGGIPGVGHLYSTHAALALRGHFLS-----SDKMCV 171
QY 127 YCCGLRSFRELTYQYRQNPASE 149
DB 172 YLLKTKAIVNASEMDIQNVPLSE 194

RESULT 9

US-10-424-599-184701
; Sequence 184701, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic David K
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 184701
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: unsure
; LOCATION: (1)-(202)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_137800C.1.pap
US-10-424-599-184701

Query Match 8.9%; Score 83.5; DB 12; Length 202;
Best Local Similarity 20.1%; Pred. No. 1.2;
Matches 37; Conservative 19; Mismatches 65; Indels 63; Gaps 7;

QY 4 QGSHALCTCCFQPMPPRRAREQDPRVAPQCAVCLQPFCHLYWGCTRGCGYCLAPFCE 63
DB 42 QWSSGICACACDD-----MQSCMLLPYPTSFXY-----KGCIGCLCP--- 78
QY 64 LNLGDKCLDGLVNNNSYESDILKNYLATRGLTWKNMLTESLVALQGVFLLSDYRVGTDT 123
DB 79 -----CFLPGKADFLGSGTFLGSCVTHFILM-SVNTACCLTDGLF-----WGLPGCL 127
QY 124 VLYCCGLRSFRELTY-----QYRO-----NIPASELPVAV 154
DB 128 VSCVACGYKRALSKYNLPAQCGDFVTHFCCHPCAICQYREIRSRSGDCEATDLKLAV 187
QY 155 TSRR 158
DB 188 VTAP 191

RESULT 10

US-10-369-493-5166
; Sequence 5166, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5166
; LENGTH: 2824
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5166

Query Match 8.9%; Score 83.5; DB 15; Length 2824;
Best Local Similarity 27.1%; Pred. No. 34;
Matches 55; Conservative 12; Mismatches 51; Indels 85; Gaps 15;

QY 10 CTCGCFQ-----MPDRAREQDPRVAPQCAVCLQ-----P 41
DB 1407 CLCNPPPCPSTELLPOR--DQIGDTLNSCQNGVCLQMGMDCAECNEFIVDRENGTP 1464
QY 42 FCH-----LYMGCTRTGCGC-----LAPFCELN-LGDKC-LDGLVNNNSYE---SD 83
DB 1465 FCKPNHCLFTQNRNSGYDCSTGRENALFLCP--NLNEFGDFCOYEGOLYNNSTYILSA 1522
QY 84 ILKNYLATRGLTWKNMLTESLVALQGVFLLSDYRVGTDTVLCVCCGLRSFRELTYQYRQ 143
DB 1523 GIENAIATN-----LCESLDTL-----YAVPN-----TFCV-----Q 1549
QY 144 NIPASELPVAVTSRPPDCVWGRNC 166
DB 1550 N-PTSTTP-SIHRCDPCYGGENC 1570

RESULT 11

US-10-424-599-224914
; Sequence 224914, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224914
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_45128C.1.pap
US-10-424-599-224914

Query Match 8.7%; Score 82; DB 12; Length 129;
Best Local Similarity 19.1%; Pred. No. 0.97;
Matches 22; Conservative 26; Mismatches 31; Indels 36; Gaps 4;

QY 15 QPMDRAREQDPRVAPQCAVCLQPF-----CHLYWGCTRTGCGY 56
DB 22 KPLKEKQRENDKSK-----SCAICLEDFPDNEEYMLTFCNHFHEDCIVPLTSGQCPV 77
QY 57 CLAPFCELNLGDKLDGLVNNNSYESDILKNYLATRGLTWKNMLTESLVALQGV 111
DB 78 CRFVICEIGRG-----NHSSFNNDIAN-----LEPSNLINGELLSILRAM 118

;; PRIOR APPLICATION NUMBER: 60/228,432
;; PRIOR FILING DATE: 2000-08-28
;; NUMBER OF SEQ ID NOS: 744
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 728
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Unknown Organism
;; FEATURE:
;; OTHER INFORMATION: Description of Unknown Organism: 85P1B3/OIPS
;; OTHER INFORMATION: clone A protein
US-09-942-052-728

Query Match 9.0%; Score 85; DB 10; Length 229;
Best Local Similarity 26.6%; Pred. No. 0.96;
Matches 38; Conservative 15; Mismatches 52; Indels 38; Gaps 6;

QY 26 QDPRVAPQCAVC---LQPFCHLYWGCTRTGCGYCLAPFCELNGLDKLDGVLNNSYES 82
Db 71 QPERCAVFQCAQCHAVLADSVHLAWDLR-----SLGAVVFSRVTTNNVLEA 117

QY 83 DILKNYLAT-RGLTWKML-----TESLVALQRGVFLLSDYRVGTDTVLC 126
Db 118 PFLVIGIEGSLKSTY-NLLFCGSGIPVGFPHLYSTHAALAAALRGHFCLS-----SDKMVC 171

QY 127 YCCGLRSFRELTYQYRONIPASE 149
Db 172 YLLTKAIVNASEMDIQNVPLSE 194

RESULT 6
US-09-942-052-729
;; Sequence 729, Application US/09942052
;; Publication No. US20030170626A1
;; GENERAL INFORMATION:
;; APPLICANT: Raitano, Arthur B.
;; APPLICANT: Paris, Mary
;; APPLICANT: Hubert, Rene S.
;; APPLICANT: Afar, Daniel
;; APPLICANT: Ge, Wangmao
;; APPLICANT: Challita-Eid, Pia M.
;; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
;; FILE REFERENCE: 51158-20028.00
;; CURRENT APPLICATION NUMBER: US/09/942,052
;; CURRENT FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: 60/228,432
;; PRIOR FILING DATE: 2000-08-28
;; NUMBER OF SEQ ID NOS: 744
;; SOFTWARE: PatentIn ver. 2.1
;; SEQ ID NO 729
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Unknown Organism
;; FEATURE:
;; OTHER INFORMATION: Description of Unknown Organism: 85P1B3/OIPS
;; OTHER INFORMATION: protein sequence
US-09-942-052-729

Query Match 9.0%; Score 85; DB 10; Length 229;
Best Local Similarity 26.6%; Pred. No. 0.96;
Matches 38; Conservative 15; Mismatches 52; Indels 38; Gaps 6;

QY 26 QDPRVAPQCAVC---LQPFCHLYWGCTRTGCGYCLAPFCELNGLDKLDGVLNNSYES 82
Db 71 QPERCAVFQCAQCHAVLADSVHLAWDLR-----SLGAVVFSRVTTNNVLEA 117

QY 83 DILKNYLAT-RGLTWKML-----TESLVALQRGVFLLSDYRVGTDTVLC 126
Db 118 PFLVIGIEGSLKSTY-NLLFCGSGIPVGFPHLYSTHAALAAALRGHFCLS-----SDKMVC 171

QY 127 YCCGLRSFRELTYQYRONIPASE 149

Db 172 YLLTKAIVNASEMDIQNVPLSE 194

RESULT 7
US-09-942-052-730
;; Sequence 730, Application US/09942052
;; Publication No. US20030170626A1
;; GENERAL INFORMATION:
;; APPLICANT: Raitano, Arthur B.
;; APPLICANT: Paris, Mary
;; APPLICANT: Hubert, Rene S.
;; APPLICANT: Afar, Daniel
;; APPLICANT: Ge, Wangmao
;; APPLICANT: Challita-Eid, Pia M.
;; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
;; FILE REFERENCE: 51158-20028.00
;; CURRENT APPLICATION NUMBER: US/09/942,052
;; CURRENT FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: 60/228,432
;; PRIOR FILING DATE: 2000-08-28
;; NUMBER OF SEQ ID NOS: 744
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 730
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Unknown Organism
;; FEATURE:
;; OTHER INFORMATION: Description of Unknown Organism: 85P1B3 protein
;; OTHER INFORMATION: sequence
US-09-942-052-730

Query Match 9.0%; Score 85; DB 10; Length 229;
Best Local Similarity 26.6%; Pred. No. 0.96;
Matches 38; Conservative 15; Mismatches 52; Indels 38; Gaps 6;

QY 26 QDPRVAPQCAVC---LQPFCHLYWGCTRTGCGYCLAPFCELNGLDKLDGVLNNSYES 82
Db 71 QPERCAVFQCAQCHAVLADSVHLAWDLR-----SLGAVVFSRVTTNNVLEA 117

QY 83 DILKNYLAT-RGLTWKML-----TESLVALQRGVFLLSDYRVGTDTVLC 126
Db 118 PFLVIGIEGSLKSTY-NLLFCGSGIPVGFPHLYSTHAALAAALRGHFCLS-----SDKMVC 171

QY 127 YCCGLRSFRELTYQYRONIPASE 149
Db 172 YLLTKAIVNASEMDIQNVPLSE 194

RESULT 8
US-09-942-052-731
;; Sequence 731, Application US/09942052
;; Publication No. US20030170626A1
;; GENERAL INFORMATION:
;; APPLICANT: Raitano, Arthur B.
;; APPLICANT: Paris, Mary
;; APPLICANT: Hubert, Rene S.
;; APPLICANT: Afar, Daniel
;; APPLICANT: Ge, Wangmao
;; APPLICANT: Challita-Eid, Pia M.
;; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
;; FILE REFERENCE: 51158-20028.00
;; CURRENT APPLICATION NUMBER: US/09/942,052
;; CURRENT FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: 60/228,432
;; PRIOR FILING DATE: 2000-08-28
;; NUMBER OF SEQ ID NOS: 744
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 731
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Homo sapiens

Db 596 GDTVLVCCGLRSFRELTYQYRQNIASELPVAVTSRPPDCYWGNC 641

RESULT 2

US-10-094-749-1799
; Sequence 1799, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1799
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-1799

Query Match 46.4%; Score 438; DB 15; Length 128;
Best Local Similarity 98.8%; Pred. No. 6.7e-39;
Matches 80; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 86 KNYLATRGLTWKXMLTESLVALQGVFLSDYRVGTDLVLCYCCGLRSFRELTYQYRQNI 145
Db 25 QNYLATRGLTWKXMLTESLVALQGVFLSDYRVGTDLVLCYCCGLRSFRELTYQYRQNI 84
Qy 146 PASELPVAVTSRPPDCYWGNC 166
Db 85 PASELPVAVTSRPPDCYWGNC 105

RESULT 3

US-10-108-260A-4516
; Sequence 4516, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5459
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4516
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4516

Query Match 39.9%; Score 376; DB 15; Length 349;
Best Local Similarity 98.4%; Pred. No. 1.1e-31;
Matches 61; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CPLQGSALCTCCQPMPPDRAREQDPRVAPQCAVCLQPFCHLYWGCTRTGCVGCLAP 60
Db 276 CPLQGSALCTCCQPMPPDRAREQDPRVAPQCAVCLQPFCHLYWGCTRTGCVGCLAP 335
Qy 61 FC 62
Db 336 FC 337
RESULT 4
US-10-424-599-264601
; Sequence 264601, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 264601
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80956C.1.pap
US-10-424-599-264601

Query Match 16.0%; Score 150.5; DB 12; Length 287;
Best Local Similarity 25.1%; Pred. No. 1.2e-07;
Matches 46; Conservative 26; Mismatches 84; Indels 27; Gaps 6;
Qy 1 CPLQGSALCTCCQPMPPDRAREQDPRVAPQCAVCLQPFCHLYW---GCTRTGCVGC 57
Db 90 CRYDTVHLQCAQCGGMPSTGFG-----IPQYCGCDRSPCGAYWALGVGTNGSY-- 141
Qy 58 LAPFCELN---LGGKLDGV---LNNNSVESDILKNYLATRGLTWKXMLTESLVALQ- 108
Db 142 --FVCSQDTRLRISHSISRIPLLAHEKXNLHEQNITDSCIRQMGRTLPDVISWEIAFEN 199
Qy 109 -----RGVFLSDYRVGTDLVLCYCCGLRSFRELTYQYRQNIASELPVAVTSRPPDCYWG 163
Db 200 REIDRRMWMLEAMITARTFVCQDCTYKLVSFLLYWFRLSIPKHLPPDESAREDCWYG 259
Qy 164 RNC 166
Db 260 YAC 262

RESULT 5

US-09-942-052-728
; Sequence 728, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eld, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:49:13 ; Search time 26.2936 Seconds
(without alignments)
1752.371 Million cell updates/sec

Title: US-10-048-046-2_COPY_476_641

Perfect score: 943

Sequence: 1 CFIQSGSHALCTCCFQMPDR.....ASELPVATSRPCYWGRCNC 166

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	943	100.0	664	9	US-09-780-525-2 Sequence 2, Appli
2	438	46.4	128	15	US-10-094-749-1799 Sequence 1799, Ap
3	376	39.9	349	15	US-10-108-260A-4516 Sequence 4516, Ap
4	150.5	16.0	287	12	US-10-424-599-264601 Sequence 264601,
5	85	9.0	229	10	US-09-942-052-728 Sequence 728, App
6	85	9.0	229	10	US-09-942-052-729 Sequence 729, App
7	85	9.0	229	10	US-09-942-052-730 Sequence 730, App
8	85	9.0	229	10	US-09-942-052-731 Sequence 731, App
9	83.5	8.9	202	12	US-10-424-599-184701 Sequence 184701,
10	83.5	8.9	2824	15	US-10-369-493-5166 Sequence 5166, Ap
11	82	8.7	129	12	US-10-424-599-224914 Sequence 224914,
12	78.5	8.3	216	12	US-10-425-114-62319 Sequence 62319, A
13	77	8.2	303	15	US-10-108-260A-3965 Sequence 3965, Ap
14	77	8.2	771	12	US-10-142-426-495 Sequence 495, App
15	77	8.2	771	14	US-10-123-155-495 Sequence 495, App

16	77	8.2	771	14	US-10-146-731-495 Sequence 495, App
17	77	8.2	771	14	US-10-140-472-495 Sequence 495, App
18	77	8.2	771	14	US-10-141-761-495 Sequence 495, App
19	77	8.2	771	14	US-10-142-885-495 Sequence 495, App
20	77	8.2	771	14	US-10-158-790-495 Sequence 495, App
21	77	8.2	771	15	US-10-137-871-495 Sequence 495, App
22	77	8.2	771	15	US-10-140-923-495 Sequence 495, App
23	77	8.2	771	15	US-10-141-756-495 Sequence 495, App
24	77	8.2	771	15	US-10-141-759-495 Sequence 495, App
25	77	8.2	771	15	US-10-140-805-495 Sequence 495, App
26	77	8.2	771	15	US-10-140-864-495 Sequence 495, App
27	76.5	8.1	487	15	US-10-369-493-22178 Sequence 22178, A
28	76	8.1	567	12	US-10-369-022-2 Sequence 2, Appli
29	75.5	8.0	419	14	US-10-354-358-76 Sequence 76, Appli
30	75.5	8.0	615	10	US-03-858-909-2 Sequence 2, Appli
31	75.5	8.0	615	14	US-10-172-712-30 Sequence 30, Appli
32	75.5	8.0	615	16	US-10-449-132-2 Sequence 2, Appli
33	75	8.0	1835	10	US-09-457-571-15 Sequence 15, Appli
34	74.5	7.9	246	15	US-10-104-047-3660 Sequence 3660, Ap
35	74.5	7.9	2849	12	US-10-142-426-371 Sequence 371, App
36	74.5	7.9	2849	14	US-10-123-155-371 Sequence 371, App
37	74.5	7.9	2849	14	US-10-146-731-371 Sequence 371, App
38	74.5	7.9	2849	14	US-10-140-472-371 Sequence 371, App
39	74.5	7.9	2849	14	US-10-141-761-371 Sequence 371, App
40	74.5	7.9	2849	14	US-10-142-885-371 Sequence 371, App
41	74.5	7.9	2849	14	US-10-158-790-371 Sequence 371, App
42	74.5	7.9	2849	15	US-10-137-871-371 Sequence 371, App
43	74.5	7.9	2849	15	US-10-140-923-371 Sequence 371, App
44	74.5	7.9	2849	15	US-10-141-756-371 Sequence 371, App
45	74.5	7.9	2849	15	US-10-141-759-371 Sequence 371, App

ALIGNMENTS

RESULT 1

US-09-780-525-2
; Sequence 2, Application US/09780525
; Patent No. US20020004223A1
; GENERAL INFORMATION:
; APPLICANT: Bin-Bing Zhou
; APPLICANT: Yuan Zhu
; APPLICANT: Priya Chaturvedi
; APPLICANT: Mark R. Huxle
; APPLICANT: Xiaotong Li
; TITLE OF INVENTION: FHAR1, A NEW RING FINGER PROTEIN
; FILE REFERENCE: GP-70668-C1
; CURRENT APPLICATION NUMBER: US/09/780,525
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 09/456,876
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 664
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-780-525-2

Query Match	100.0%;	Score	943;	DB	9;	Length	664;
Best Local Similarity	100.0%;	Pred. No.	4e-92;				
Matches	166;	Conservative	0;	Mismatches	0;	Indels	0;
QY	1	CFLQSGSHALCTCCFQMPDRRAEREQDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP	60				
Db	476	CFLQSGSHALCTCCFQMPDRRAEREQDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP	535				
QY	61	FCELNLGDKCLDGLVNNNSYESDILKNYLATRLGLTWKNNLTSLVALQGVFLLSDYRVT	120				
Db	536	FCELNLGDKCLDGLVNNNSYESDILKNYLATRLGLTWKNNLTSLVALQGVFLLSDYRVT	595				
QY	121	GDTVLCVCCGLRSFRLTYQYRONIPASELPVATSRDDCTYWGRCNC	166				

RESULT 15
US-09-798-051-9
; Sequence 9, Application US/09798051
; Patent No. 6632780
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ke
; APPLICANT: Cam, Linh
; APPLICANT: Nakayama, Naoki
; TITLE OF INVENTION: Chordin-Like-2 Molecules and Uses Thereof
; FILE REFERENCE: 01-005
; CURRENT APPLICATION NUMBER: US/09/798,051
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-051-9

Query Match 7.2%; Score 67.5; DB 4; Length 452;
Best Local Similarity 25.0%; Pred. No. 27;
Matches 40; Conservative 18; Mismatches 73; Indels 29; Gaps 9;

QY	4	QGSHALCTCCFQMPDRAEREQDPVAPQCCAVCLQPECHLYWG--CTRITGC--YGCLIA	59
DB	23	QVKHSETYCMFQDKKYRVGER-WHPYLEPYGLVYCVNCICSENGNVLCSEVRCPNVHCLS	81
QY	60	P-----FCEINLGDKCLDGVINNNSYESDILKNYLAIRGLTWKNMLTESLVALQRGVFL	114
DB	82	PVHIPHLCCPCPDSLPVWNKVTSKS-----CEYNGTTYQH--GELFVA--EGLF--	129
QY	115	SDYRVGTGDTVLCYC-----CGLRSPRELTYYQRNIPAS	148
DB	130	-QNRQPNQCTQCSCSEGNVYGLKTCFKLTCAFPVSVSDS	168

Search completed: May 7, 2004, 14:51:24
Job time : 12.0433 secs

```
; Sequence 5, Application US/09798051
; Patent No. 6632780
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ke
; APPLICANT: Cam, Linh
; APPLICANT: Nakayama, Naoki
; TITLE OF INVENTION: Chordin-Like-2 Molecules and Uses Thereof
; FILE REFERENCE: 01-005
; CURRENT APPLICATION NUMBER: US/09/798,051
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 5
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-051-5

Query Match
Best Local Similarity 24.3%; Pred. No. 22; Length 429;
Matches 25; Conservative 8; Mismatches 46; Indels 24; Gaps 3;

Qy 2 PLOGSHALCTCCFOP-----MPDRRAE-----REQDPRVAPQCCAVCLQ 40
Db 84 PVTEPQCCCKVEPHPTPSGLRAPPKSCQNGTMYQHGEIFSAHELFFSLPQNCVLCSC 143
Qy 41 PFCHLYWGCTRTGCGCLAPFCEHLNLDGKLDGLVNNNSYED 83
Db 144 TEGQIYGLTTCPEFGPAP---LPLPDSGCCQACKDEASEQSD 183

RESULT 13
US-08-191-866D-81
; Sequence 81, Application US/08191866D
; Patent No. 5783195
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
; APPLICANT: Macdonald, Richard D.
; TITLE OF INVENTION: Recombinant Infectious Bovine
; TITLE OF INVENTION: Rhinotracheitis Virus S-1ER-052 And Uses Thereof
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/191,866D
; FILING DATE: 4 February 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 572 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-191-866D-81

Query Match
7.2%; Score 68; DB 1; Length 572;
Best Local Similarity 20.4%; Pred. No. 32;
Matches 31; Conservative 14; Mismatches 41; Indels 66; Gaps 5;

Qy 49 CTRTGCYGCCLAPFC-----ELNLGDKLDGLVNNNSYEDILKN----- 87
Db 350 CNTTGCPTQDCNQASYPWFSNRRMVNSIIIVVDKGDATFSLRVVTIPMSQNYWGSE 409
Qy 88 -----YLATRLGTWKNMLTESLVALQRGVFLLSDYRTVGTVLCYCCGLSREL 137
Db 410 GRLLGLGDRIIYIYTRSTSWHS-----KLQGLVIDISDY----- 442
Qy 138 TYQYRONIPASELPVAVTSRP---DCYWGRC 166
Db 443 -----NNIRINWTHNVPSRPGNDECPWGHSC 469

RESULT 14
US-08-185-949B-81
; Sequence 81, Application US/08185949B
; Patent No. 5874279
; GENERAL INFORMATION:
; APPLICANT: Mark D. Cochran
; APPLICANT: Richard D. Macdonald
; TITLE OF INVENTION: Recombinant Infectious Bovine
; TITLE OF INVENTION: Rhinotracheitis Virus
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM 330 466 DX2
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,949B
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 572 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-185-949B-81

Query Match
7.2%; Score 68; DB 2; Length 572;
Best Local Similarity 20.4%; Pred. No. 32;
Matches 31; Conservative 14; Mismatches 41; Indels 66; Gaps 5;

Qy 49 CTRTGCYGCCLAPFC-----ELNLGDKLDGLVNNNSYEDILKN----- 87
Db 350 CNTTGCPTQDCNQASYPWFSNRRMVNSIIIVVDKGDATFSLRVVTIPMSQNYWGSE 409
Qy 88 -----YLATRLGTWKNMLTESLVALQRGVFLLSDYRTVGTVLCYCCGLSREL 137
Db 410 GRLLGLGDRIIYIYTRSTSWHS-----KLQGLVIDISDY----- 442
Qy 138 TYQYRONIPASELPVAVTSRP---DCYWGRC 166
Db 443 -----NNIRINWTHNVPSRPGNDECPWGHSC 469
```


Db 904 CHVKGSHSAVCHLETGLDC-----KENVTGQQDQCL-----HGYGL--D 943

QY 53 GCYGCCLAPFCEL--NLGDKCLD 72
: : : : :
Db 944 SGHGRCPNCNVAGSVSDGCTD 965

RESULT 8
US-09-543-681A-5344
; Sequence 5344, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5344
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5344

Query Match 7.5%; Score 71; DB 4; Length 200;
Best Local Similarity 25.3%; Pred. No. 3.6;
Matches 20; Conservative 17; Mismatches 34; Indels 8; Gaps 3;

QY 76 NNNYSIEDILKNYLATRLGKLNWLTB--SLVALQGVFLSLD-YRVGTDTVLG-----Y 127
: : : : :
Db 5 SKNPMYQDLIRGELTEAATLSRFLQDDANIEAIAQAAVLLDSFKAGKVLSCGNGGSH 64
: : : : :
QY 128 CGLRSFRELTYQRONIP 146
: : : : :
Db 65 CDAMHFAEELTGRYENRP 83
: : : : :

RESULT 9
US-09-328-352-5062
; Sequence 5062, Application US/09328352
; Patent No. 8562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5062
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5062

Query Match 7.5%; Score 71; DB 4; Length 460;
Best Local Similarity 26.1%; Pred. No. 11;
Matches 37; Conservative 22; Mismatches 57; Indels 26; Gaps 9;

QY 26 QDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAPFCELNGLDKCLDGLVNN---NSYES 82
: : : : :
Db 333 QSPK-STQPDVICTNFKYMWNSAQQLTHHTIGG-CNVQVGLMGSGTISGSTPDSYGS 390
: : : : :
QY 83 DILKNYLATRLGKLNWLTB--SLVALQGVFLSLD-YRVGTDTVL-----CYCCGLR-SFREL 137
: : : : :
Db 391 LLELTWNTTKPLATANGET-----RG--FLQD-----GDTLIMKGHCENKIRIGFGEV 437
: : : : :
QY 138 TYQYRONIPASELPVAVTSRPD 159
: : : : :
Db 438 R---NTVLPALTFDFAETSEPN 456
: : : : :

RESULT 10
US-09-845-583A-2
; Sequence 2, Application US/09845583A
; Patent No. 6635616
; GENERAL INFORMATION:
; APPLICANT: Burgeson, Robert
; APPLICANT: Brunken, William Joseph
; APPLICANT: Champiaud, Marie-France
; APPLICANT: Hunter, Dale
; TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
; FILE REFERENCE: 10287-056001
; CURRENT APPLICATION NUMBER: US/09/845,583A
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/200,863
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3635
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-845-583A-2

Query Match 7.5%; Score 71; DB 4; Length 3635;
Best Local Similarity 24.5%; Pred. No. 1.7e+02;
Matches 25; Conservative 9; Mismatches 38; Indels 30; Gaps 6;

QY 1 CPLQGSHAL-----CTCCFOPMPDRRAERQDPRVAPQCAVCLQ 41
Db 1379 CFCRG-HVIGRDCSRCATGYGFCNCRPCDCGRLCDLGTQCICPRTVPPDCLVC-QP 1436
: : : : :
QY 42 ---FCHLYWGCTRTGCVG-----CLAPFCELNGLD-KCLDGV 74
: : : : :
Db 1437 GSFCHPLVGCCEGNCSGFGVQELTDTCDMDSCQCRPNV 1478
: : : : :

RESULT 11
US-09-798-051-6
; Sequence 6, Application US/09798051
; Patent No. 6632780
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ke
; APPLICANT: Cam, Linh
; APPLICANT: Nakayama, Naoki
; TITLE OF INVENTION: Chordin-Like-2 Molecules and Uses Thereof
; FILE REFERENCE: 01-005
; CURRENT APPLICATION NUMBER: US/09/798,051
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-051-6

Query Match 7.2%; Score 68; DB 4; Length 408;
Best Local Similarity 24.3%; Pred. No. 20;
Matches 25; Conservative 8; Mismatches 46; Indels 24; Gaps 3;

QY 2 PLQGSHALCTCCFQP-----MPDRRAE-----REQDPRVAPQCAVCLQ 40
: : : : :
Db 63 PYTEFQCCPKCVBHTPSGLRAPKSCQHNGTMYQHGEIFSAHELFPRLPNCVLCSC 122
: : : : :
QY 41 PFCHLYWGCTRTGCVGCLAPFCELNGLDKCLDGLVNNNSYESD 83
: : : : :
Db 123 TEGQIYCGLTTCPEPGCAP---LPLPDSCCQACKDEASEQSD 162
: : : : :

RESULT 12
US-09-798-051-5

```
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1835 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-836-325-15

Query Match      8.0%; Score 75; DB 3; Length 1835;
Best Local Similarity 19.9%; Pred. No. 24;
Matches 29; Conservative 26; Mismatches 59; Indels 32; Gaps 6;

QY 4 QGSHALCTCCFPMPDRRAEQPRVAPQCAVCLQPFCHLYMGCTRTGCGCLAPFCE 63
Db 1023 RSSSESGSTVDNPLFGEGERAEAFNDPEACFTD-----GCVRFEC-----CQ 1066

QY 64 INLG-----DKLDGVLNNVSEIDILKRYLATRGLTWKMLTESIVALQGVFLSD 116
Db 1067 VNSGKGWIRKTCY--IVEHSWFESFIVLMILLSSGA---LAFEDIYIEKTKIKILE 1120

QY 117 YRVTGDTVLQVCCGLRSFRLTYQVR 142
Db 1121 Y---ADKIFTYIFILEMLLWAYGYK 1143

RESULT 6
US-08-460-309-5
; Sequence 5, Application US/08460309
; Patent No. 5837496
; GENERAL INFORMATION:
; APPLICANT: Engvall, Eva
; APPLICANT: Leivo, Ilmo
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
; TITLE OF INVENTION: Fragments and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460.309
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/125,077
; FILING DATE: 22-SEP-1993
; APPLICATION NUMBER: US PCT/US 94/10730
; FILING DATE: 21-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,319
; FILING DATE: 30-JAN-1990
; APPLICATION NUMBER: US 07/919,951
; FILING DATE: 27-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9721
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3075 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-125-077-5

Query Match      7.6%; Score 72; DB 2; Length 3075;
Best Local Similarity 28.0%; Pred. No. 1e+02;
Matches 23; Conservative 8; Mismatches 21; Indels 30; Gaps 5;
```

```
; TYPE: amino acid
; TOPOLOGY: linear
US-08-460-309-5

Query Match      7.6%; Score 72; DB 2; Length 3075;
Best Local Similarity 28.0%; Pred. No. 1e+02;
Matches 23; Conservative 8; Mismatches 21; Indels 30; Gaps 5;

QY 1 CPLQGSHA-----LCTCCFPMPDRRAEQPRVAPQCAVCLQPFCHLYMGCTRT 52
Db 904 CHVKGSHSACHLEFGLCDC-----KENVTGQCCDQCL-----HGYGL--D 943

QY 53 GCVGCLAPFCEL--NLGDKCLD 72
Db 944 SGHGCRPCNCSVAGSVSDGCTD 965

RESULT 7
US-08-125-077-5
; Sequence 5, Application US/08125077
; Patent No. 5872231
; Patent No. 5872231 5840863
; GENERAL INFORMATION:
; APPLICANT: Engvall, Eva
; APPLICANT: Leivo, Ilmo
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
; TITLE OF INVENTION: Fragments and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,077
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US PCT/US 94/10730
; FILING DATE: 21-SEP-1994
; APPLICATION NUMBER: US 07/472,319
; FILING DATE: 30-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/919,951
; FILING DATE: 27-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9721
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3075 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-125-077-5

Query Match      7.6%; Score 72; DB 2; Length 3075;
Best Local Similarity 28.0%; Pred. No. 1e+02;
Matches 23; Conservative 8; Mismatches 21; Indels 30; Gaps 5;
```

```
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12650

Query Match      8.5%; Score 80; DB 4; Length 197;
Best Local Similarity 30.6%; Pred. No. 0.34;
Matches 22; Conservative 15; Mismatches 27; Indels 8; Gaps 3;

QY 83 DILKNYLRGLTWKMLTE--SLVALQGVFLLS--YRVTDGTVLC-----YCCGLRSF 134
DB 8 DIRMELNBAETLANFLODEANIHAIQRAVLLADSFAGGKVLSCNGGSHCDAMEHA 67
QY 135 RELTYQYRONIP 146
DB 68 EELTGRYRENRP 79

RESULT 3
US-08-841-483-2
; Sequence 2, Application US/08841483B
; Patent No. 5976875
; GENERAL INFORMATION:
; APPLICANT: Prescott, Steven M.
; APPLICANT: Bunting, Michaeline
; APPLICANT: Tang, Wen
; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 2037.2.1a
; CURRENT APPLICATION NUMBER: US/08/841,483B
; CURRENT FILING DATE: 1997-04-22
; EARLIER APPLICATION NUMBER: 60/016,210
; EARLIER FILING DATE: 1996-04-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-841-483-2

Query Match      8.1%; Score 76; DB 2; Length 567;
Best Local Similarity 21.6%; Pred. No. 3.9;
Matches 36; Conservative 21; Mismatches 54; Indels 56; Gaps 10;

QY 3 LQSHALCTCCFQMPD---RRAER-----EODPRVAPQ-----CAVCL 39
DB 82 LOG--AFDCGGLRVDEGCLRKADRFQCKEIMLKNDTKVLDAMPHHWIRGNVPLCSYM 139
QY 40 QPFCHLYWGCTRTGC-YGCLAPCELNGLDKLDGVNNSVESDILKN-----YLAT- 91
DB 140 --VCKQCGGCPKLCYRCI--WCQKTVHDECWNLSKNEKCDGFEKKNLIIPPSYLTSI 195
QY 92 -----RGLTWKNM--LTSVLALQGVFLLSDYRV 119
DB 196 NQMRKDKKTDYEVLSKGLGKQWTPLLIILANSRSGTNNMGEGLGEFRI 242

RESULT 4
US-09-382-911-2
; Sequence 2, Application US/09382911
; Patent No. 6221658
; GENERAL INFORMATION:
; APPLICANT: Prescott, Steven M.
; APPLICANT: Bunting, Michaeline
; APPLICANT: Tang, Wen
; APPLICANT: Topham, Matthew
; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 2037.2.1a
; CURRENT APPLICATION NUMBER: US/09/382,911
; CURRENT FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 08/841,483
; PRIOR FILING DATE: 1997-04-22

; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12650

Query Match      8.5%; Score 80; DB 4; Length 197;
Best Local Similarity 30.6%; Pred. No. 0.34;
Matches 22; Conservative 15; Mismatches 27; Indels 8; Gaps 3;

QY 83 DILKNYLRGLTWKMLTE--SLVALQGVFLLS--YRVTDGTVLC-----YCCGLRSF 134
DB 8 DIRMELNBAETLANFLODEANIHAIQRAVLLADSFAGGKVLSCNGGSHCDAMEHA 67
QY 135 RELTYQYRONIP 146
DB 68 EELTGRYRENRP 79

RESULT 3
US-08-841-483-2
; Sequence 2, Application US/08841483B
; Patent No. 5976875
; GENERAL INFORMATION:
; APPLICANT: Prescott, Steven M.
; APPLICANT: Bunting, Michaeline
; APPLICANT: Tang, Wen
; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 2037.2.1a
; CURRENT APPLICATION NUMBER: US/08/841,483B
; CURRENT FILING DATE: 1997-04-22
; EARLIER APPLICATION NUMBER: 60/016,210
; EARLIER FILING DATE: 1996-04-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-841-483-2

Query Match      8.1%; Score 76; DB 2; Length 567;
Best Local Similarity 21.6%; Pred. No. 3.9;
Matches 36; Conservative 21; Mismatches 54; Indels 56; Gaps 10;

QY 3 LQSHALCTCCFQMPD---RRAER-----EODPRVAPQ-----CAVCL 39
DB 82 LOG--AFDCGGLRVDEGCLRKADRFQCKEIMLKNDTKVLDAMPHHWIRGNVPLCSYM 139
QY 40 QPFCHLYWGCTRTGC-YGCLAPCELNGLDKLDGVNNSVESDILKN-----YLAT- 91
DB 140 --VCKQCGGCPKLCYRCI--WCQKTVHDECWNLSKNEKCDGFEKKNLIIPPSYLTSI 195
QY 92 -----RGLTWKNM--LTSVLALQGVFLLSDYRV 119
DB 196 NQMRKDKKTDYEVLSKGLGKQWTPLLIILANSRSGTNNMGEGLGEFRI 242

RESULT 4
US-09-382-911-2
; Sequence 2, Application US/09382911
; Patent No. 6221658
; GENERAL INFORMATION:
; APPLICANT: Prescott, Steven M.
; APPLICANT: Bunting, Michaeline
; APPLICANT: Tang, Wen
; APPLICANT: Topham, Matthew
; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 2037.2.1a
; CURRENT APPLICATION NUMBER: US/09/382,911
; CURRENT FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 08/841,483
; PRIOR FILING DATE: 1997-04-22

; ORGANISM: Homo sapiens
US-09-382-911-2

Query Match      8.1%; Score 76; DB 3; Length 567;
Best Local Similarity 21.6%; Pred. No. 3.9;
Matches 36; Conservative 21; Mismatches 54; Indels 56; Gaps 10;

QY 3 LQSHALCTCCFQMPD---RRAER-----EODPRVAPQ-----CAVCL 39
DB 82 LOG--AFDCGGLRVDEGCLRKADRFQCKEIMLKNDTKVLDAMPHHWIRGNVPLCSYM 139
QY 40 QPFCHLYWGCTRTGC-YGCLAPCELNGLDKLDGVNNSVESDILKN-----YLAT- 91
DB 140 --VCKQCGGCPKLCYRCI--WCQKTVHDECWNLSKNEKCDGFEKKNLIIPPSYLTSI 195
QY 92 -----RGLTWKNM--LTSVLALQGVFLLSDYRV 119
DB 196 NQMRKDKKTDYEVLSKGLGKQWTPLLIILANSRSGTNNMGEGLGEFRI 242

RESULT 5
US-08-836-325-15
; Sequence 15, Application US/08836325
; Patent No. 6110672
; GENERAL INFORMATION:
; APPLICANT: Mandel, Gail
; APPLICANT: Halsegouda, Simon
; APPLICANT: Borden, Laurence A.
; TITLE OF INVENTION: Peripheral Nervous System Specific
; TITLE OF INVENTION: Sodium Channels, DNA Encoding Therefor, Crystallization,
; TITLE OF INVENTION: X-ray Diffraction, Computer Molecular Modeling, Rational
; TITLE OF INVENTION: Drug Design, Drug Screening, and Methods of Making and Using
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C
; STREET: 1100 New York Ave., N. W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,325
; FILING DATE: 2-MAY-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14251
; FILING DATE: 02-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/482,401
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/334,029
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0917.0240002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:42:23 ; Search time 11.0433 Seconds
(without alignments)
776.028 Million cell updates/sec

Title: US-10-048-046-2_COPY_476_641
Perfect score: 943
Sequence: 1 CPGQSHALCTCCFQMPDR.....ASELPVAVTSRDPICYGRNC 166
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	8.6	376	4	US-09-489-039A-11343
2	80	8.5	197	4	US-09-489-039A-12650
3	76	8.1	567	2	US-08-841-483-2
4	76	8.1	567	3	US-09-382-911-2
5	75	8.0	1835	3	US-08-836-325-15
6	72	7.6	3075	2	US-08-460-309-5
7	72	7.6	3075	2	US-08-125-077-5
8	71	7.5	200	4	US-09-543-681A-5344
9	71	7.5	460	4	US-09-328-352-5062
10	71	7.5	3635	4	US-09-843-583A-2
11	68	7.2	408	4	US-09-798-051-6
12	68	7.2	429	4	US-09-798-051-5
13	68	7.2	572	1	US-08-191-866D-81
14	68	7.2	572	2	US-08-185-949B-81
15	67.5	7.2	452	4	US-09-798-051-9
16	67	7.1	174	3	US-09-383-586-12
17	67	7.1	202	4	US-09-489-039A-14068
18	67	7.1	276	2	US-09-786-606-5
19	67	7.1	276	4	US-09-205-258-1027
20	67	7.1	341	4	US-09-205-258-1034
21	67	7.1	441	4	US-08-914-375C-66
22	67	7.1	449	3	US-09-306-593-10
23	67	7.1	503	3	US-09-382-256-10
24	67	7.1	503	3	US-09-395-115-10
25	67	7.1	503	4	US-08-436-265-10
26	67	7.1	503	4	US-09-679-187-10
27	67	7.1	3174	2	US-08-477-451-3

28 66.5 7.1 259 4 US-09-252-991A-29086 Sequence 29086, A
29 66.5 7.1 387 1 US-08-123-161A-12 Sequence 12, Appl
30 66.5 7.1 387 1 US-08-483-278-12 Sequence 12, Appl
31 66 7.0 125 4 US-09-489-039A-10681 Sequence 10681, A
32 66 7.0 330 4 US-09-489-039A-8234 Sequence 8234, Ap
33 66 7.0 515 4 US-09-635-872A-6 Sequence 6, Appl
34 66 7.0 515 4 US-09-636-077A-6 Sequence 6, Appl
35 66 7.0 515 4 US-09-636-060C-6 Sequence 6, Appl
36 66 7.0 515 4 US-09-985-552-6 Sequence 6, Appl
37 65.5 6.9 276 4 US-09-719-402A-782 Sequence 782, App
38 65.5 6.9 324 4 US-09-719-402A-2 Sequence 2, Appl
39 65.5 6.9 539 4 US-09-900-230-3 Sequence 3, Appl
40 65.5 6.9 1417 4 US-08-900-230-3 Sequence 28511, A
41 65 6.9 177 4 US-09-252-991A-28511 Sequence 5215, Ap
42 65 6.9 212 1 US-09-107-532A-5215 Sequence 1, Appl
43 65 6.9 1015 3 US-08-537-210A-1 Sequence 1, Appl
44 65 6.9 1015 3 US-09-113-825-1 Sequence 2, Appl
45 65 6.9 1165 1 US-08-144-121-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-489-039A-11343
; Sequence 11343, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11343
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11343

Query Match 8.6%; Score 81; DB 4; Length 376;
Best Local Similarity 33.7%; Pred. No. 0.62; 32; Indels 18; Gaps 5;
Matches 30; Conservative
Qy 32 PQCAVCL-----QPFCILYWGCTTGCYGCILAPFCELNGLGDKLDGVNNNSYS 82
Db 237 PIAAACTADAVSSTAGKPLKQLFW-C--AGSNGTLPFGNGNGK--GVIRDSSLLS 290
Qy 83 DILKNVLTATGLTWKXMTESLVALQGV 111
Db 291 ARVLAALHRRGLAWKTGSE---AMCRGV 316

RESULT 2
US-09-489-039A-12650
; Sequence 12650, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12650
; LENGTH: 197
; TYPE: PRT

PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 28-OCT-1999; 99US-0162142P.
PR 28-OCT-1999; 99US-0162142P.
Query Match 14.6%; Score 137.5; DB 3; Length 453;
Best Local Similarity 25.0%; Pred. No. 4.9e-05;
Matches 45; Conservative 26; Mismatches 88; Indels 21; Gaps 6;
QY 1 CPQGSALCTCCFQMPDRARERQDPVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 60
Db 255 CEHGAHLOCHLCQGMPPFRANLQ-----VPLHCKGCDRPPCGATWS-SENVTQGVSP 307
QY 61 FCSELN-----LGDKCLGV-----LNNNSYESDILKNYLATRGLTWKMLTESL-----VAL 107
Db 308 VCVRERFRPISERTITRIPITHEMNREHEDITQRCIAHMEKTPDVAEMLRLENNREI 367
QY 108 QRGVFLLSYR-VGTGTVLCYCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGRC 166
Db 368 DSRMPLNHAETITASTHVCNDCYDKLVGLVYFRITLPRNHLPADVAAREDCMYGYAC 427
RESULT 14
ID AAB34759 standard; protein; 227 AA.
XX AAB34759;
XX
XX 26-JAN-2001 (first entry)
XX
XX Human secreted protein fragment encoded by DNA clone vo3 1.
XX
XX Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;
XX systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;
XX haematopoiesis regulation; tissue regrowth; wound healing; haemophilia;
XX Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;
XX contraceptive; infection; growth inhibition; hyperproliferative disorder;
XX psoriasis.
XX
XX Homo sapiens.
XX
XX WO200055375-A1.
XX
XX 21-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-US007285.
XX
XX 17-MAR-1999; 99US-0124808P.
XX 17-MAR-1999; 99US-0124916P.
XX 17-AUG-1999; 99US-0149639P.
XX 01-OCT-1999; 99US-0157247P.
XX 29-NOV-1999; 99US-0167824P.
XX 15-FEB-2000; 2000US-0182711P.
XX
XX (ALPH-) ALPHAGENE INC.
XX
XX Valenzuela D, Yuan O, Hoffman E, Hall J, Rapiejko P;
XX
XX

XX WPI; 2000-639211/61.
DR N-PSDB; AAC59810.
XX
XX Novel proteins and polypeptides useful for the treatment of e.g multiple
XX sclerosis, systemic lupus erythematosus, rheumatoid arthritis, cancer,
XX Alzheimer's disease, Parkinson's disease, stroke, anemia and ulcers.
XX
XX Disclosure; Page 480-481; 493pp; English.
XX
XX This invention relates to 59 human secreted proteins and the nucleotide
XX sequences encoding them. Sequences AAC59788-C59846 and AAB34687-834745
XX represent the proteins and their encoding nucleotide sequences, and
XX sequences AAB34746-B34771 represent fragments of the proteins. Probes for
XX the DNA sequences are represented by sequences AAC59847-C59596. The
XX proteins exhibit neuroprotective, dermatological, immunosuppressive,
XX antiinflammatory, antianaemic, neurotropic, antiparkinsonian,
XX cerebroprotective, haemostatic, vulnerary, cytostatic, antipsoriatic,
XX antibacterial, virucide, and fungicide activity. The proteins and
XX nucleotide sequences are useful as nutritional sources or supplements and
XX in research. The proteins are useful for treating immune deficiency and
XX disorders, which may be genetic or resulting from infections, autoimmune
XX disorders such as multiple sclerosis, systemic lupus erythematosus,
XX rheumatoid arthritis, and for treating myeloid or lymphoid cell
XX deficiencies such as anaemias by regulating haematopoiesis. The proteins
XX are also useful in compositions for bone, cartilage, tendon, ligament
XX and/or nerve tissue growth or regeneration, for wound healing, tissue
XX repair and replacement and in the treatment of wounds, incisions and
XX ulcers. Other uses include in the treatment of central and peripheral
XX nervous system and neuropathies such as Alzheimer's and Parkinson's
XX diseases and Shy-Drager syndrome, and mechanical and traumatic disorders,
XX such as spinal cord disorders, head trauma and stroke. The proteins may
XX also be used as a contraceptive, and for treating coagulation disorders
XX such as haemophilias. The protein and nucleotide sequences with cadherin
XX activity are useful for treating cancer. Other uses for the protein
XX include for inhibiting the growth, infection or function of, or killing,
XX infectious agents such as bacteria, virus, fungi and other parasites, for
XX effecting bodily characteristics such as height, weight, hair colour,
XX effecting biorhythms or cardiac cycles or rhythms, effecting metabolism,
XX catabolism, anabolism, processing, utilization, storage or elimination of
XX dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors,
XX effecting behavioural characteristics, providing analgesic effects and
XX for treating hyperproliferative disorders such as psoriasis
XX
XX Sequence 227 AA;
XX
Query Match 9.8%; Score 92.5; DB 3; Length 227;
Best Local Similarity 23.8%; Pred. No. 0.62;
Matches 45; Conservative 21; Mismatches 70; Indels 53; Gaps 10;
QY 2 PLQGSALCTCCFQMP---DRARERQDPVAP-----QQCAVC----- 38
Db 19 PVLPSAL---LFPGLPAGFSRRLSSNAGPELLAWLACFLRLPACLLSLVALPGCWA 75
QY 39 ---LQPFCHLYWGCTRTG---CYGCLAP-FCENLGDKCLDGLVNNNSYESDILKNYLAT 91
Db 76 LSGRLLPVCFWNLCLGAPAFSCGLLPVYCHLQSGLLRPTLH-----LAPPWL-- 126
QY 92 RGLTWKNNLTESLVALQGVFLLSYRVTGDTVL---CYCGLSRFRELTYQYQNTIPAS 148
Db 127 --LAMPNLAFCAMLELELLFFRGNGRVESGKLAPKCCCGFFAFS-----KDALPGP 178
QY 149 ELPVAVTSR 157
Db 179 KIQTAVLSK 187
RESULT 15
AAU56425
ID AAU56425 standard; protein; 244 AA.
XX AAU56425;
XX

PR	28-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145088P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145192P.
PR	06-MAY-1999;	99US-0132487P.	PR	23-JUL-1999;	99US-0145145P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145218P.
PR	11-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146388P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	03-AUG-1999;	99US-0147038P.
PR	25-MAY-1999;	99US-0136021P.	PR	04-AUG-1999;	99US-0147204P.
PR	27-MAY-1999;	99US-0136332P.	PR	04-AUG-1999;	99US-0147302P.
PR	28-MAY-1999;	99US-0136782P.	PR	05-AUG-1999;	99US-0147192P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147260P.
PR	03-JUN-1999;	99US-0137528P.	PR	06-AUG-1999;	99US-0147303P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147418P.
PR	07-JUN-1999;	99US-0137724P.	PR	09-AUG-1999;	99US-0147493P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138540P.	PR	10-AUG-1999;	99US-0148171P.
PR	10-JUN-1999;	99US-0138847P.	PR	11-AUG-1999;	99US-0148319P.
PR	14-JUN-1999;	99US-0139119P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148655P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148684P.
PR	17-JUN-1999;	99US-0139492P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139454P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139455P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	18-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	24-JUN-1999;	99US-0140695P.	PR	01-SEP-1999;	99US-0151930P.
PR	28-JUN-1999;	99US-0140823P.	PR	07-SEP-1999;	99US-0152363P.
PR	29-JUN-1999;	99US-0140991P.	PR	10-SEP-1999;	99US-0153070P.
PR	30-JUN-1999;	99US-0141287P.	PR	13-SEP-1999;	99US-0153758P.
PR	01-JUL-1999;	99US-0141842P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0142154P.	PR	16-SEP-1999;	99US-0154039P.
PR	02-JUL-1999;	99US-0142055P.	PR	20-SEP-1999;	99US-0154779P.
PR	06-JUL-1999;	99US-0142390P.	PR	22-SEP-1999;	99US-0155139P.
PR	08-JUL-1999;	99US-0142803P.	PR	23-SEP-1999;	99US-0155486P.
PR	09-JUL-1999;	99US-0142920P.	PR	24-SEP-1999;	99US-0155659P.
PR	12-JUL-1999;	99US-0142977P.	PR	28-SEP-1999;	99US-0156458P.
PR	13-JUL-1999;	99US-0143542P.	PR	29-SEP-1999;	99US-0156596P.
PR	14-JUL-1999;	99US-0143624P.	PR	04-OCT-1999;	99US-0157117P.
PR	15-JUL-1999;	99US-0144005P.	PR	05-OCT-1999;	99US-0157753P.
PR	16-JUL-1999;	99US-0144085P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144332P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144333P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144335P.	PR	13-OCT-1999;	99US-0159295P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144634P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159331P.
PR			PR	14-OCT-1999;	99US-0159637P.
PR			PR	14-OCT-1999;	99US-0159638P.

PR	21-OCT-1999;	99US-0145086P.	99US-0160741P.
PR	21-OCT-1999;	99US-0145088P.	99US-0160767P.
PR	21-OCT-1999;	99US-0145089P.	99US-0160768P.
PR	21-OCT-1999;	99US-0145087P.	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	99US-0160814P.
PR	22-JUL-1999;	99US-0145088P.	99US-0160815P.
PR	22-JUL-1999;	99US-0145192P.	99US-0160980P.
PR	23-JUL-1999;	99US-0145145P.	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	99US-0160989P.
PR	23-JUL-1999;	99US-0145276P.	99US-0161404P.
PR	26-JUL-1999;	99US-0145913P.	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	99US-0161360P.
PR	28-AUG-1999;	99US-0146386P.	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	
PR	05-AUG-1999;	99US-0147192P.	
PR	05-AUG-1999;	99US-0147260P.	
PR	06-AUG-1999;	99US-0147303P.	
PR	06-AUG-1999;	99US-0147416P.	
PR	08-AUG-1999;	99US-0147493P.	
PR	09-AUG-1999;	99US-0147935P.	
PR	10-AUG-1999;	99US-0148171P.	
PR	11-AUG-1999;	99US-0148319P.	
PR	12-AUG-1999;	99US-0148341P.	
PR	13-AUG-1999;	99US-0148565P.	
PR	13-AUG-1999;	99US-0148684P.	
PR	16-AUG-1999;	99US-0149368P.	
PR	17-AUG-1999;	99US-0149175P.	
PR	18-AUG-1999;	99US-0149426P.	
PR	20-AUG-1999;	99US-0149722P.	
PR	20-AUG-1999;	99US-0149723P.	
PR	20-AUG-1999;	99US-0149929P.	
PR	20-AUG-1999;	99US-0149802P.	
PR	23-AUG-1999;	99US-0149930P.	
PR	23-AUG-1999;	99US-0150566P.	
PR	25-AUG-1999;	99US-0150884P.	
PR	26-AUG-1999;	99US-0151065P.	
PR	27-AUG-1999;	99US-0151066P.	
PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153070P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157732P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159234P.	
PR	13-OCT-1999;	99US-0159255P.	
PR	14-OCT-1999;	99US-0159349P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	18-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159534P.	

Query Match 14.6%; Score 137.5; DB 3; Length 453;
 Best Local Similarity 25.0%; Pred. No. 4.9e-05;
 Matches 45; Conservative 26; Mismatches 88; Indels 21; Gaps 6;

QY	1	CPLQSSHALCTCCFQPMFDRRAEREQDPRAVAPQCAVCLQPFCHLYWGCTTGCYGC	60
DB	255	CEHHGAHLQCHLCQGMPPFRANLQ-----VPLHCKGCDRPFCCGAYNS-SENVTQGVSGP	307
QY	61	FCELN-----LGDKCLDGV-----LNNNSYESDILKNYLATRGLTWKNMLTESL----	107
DB	308	VCVRETFREISERTITRIPFITHENRHEQDITQRCIAHMEKTPDVAEMLRLENNREI	367</

PR	21-OCT-1999;	99US-0160767P.	PR	30-APR-1999;	99US-0132048P.
PR	21-OCT-1999;	99US-0160768P.	PR	30-APR-1999;	99US-0132407P.
PR	21-OCT-1999;	99US-0160770P.	PR	04-MAY-1999;	99US-0132484P.
PR	21-OCT-1999;	99US-0160814P.	PR	05-MAY-1999;	99US-0132485P.
PR	21-OCT-1999;	99US-0160815P.	PR	06-MAY-1999;	99US-0132486P.
PR	22-OCT-1999;	99US-0160980P.	PR	06-MAY-1999;	99US-0132487P.
PR	22-OCT-1999;	99US-0160981P.	PR	07-MAY-1999;	99US-0132862P.
PR	22-OCT-1999;	99US-0160989P.	PR	11-MAY-1999;	99US-0134258P.
PR	22-OCT-1999;	99US-0161404P.	PR	14-MAY-1999;	99US-0134218P.
PR	25-OCT-1999;	99US-0161405P.	PR	14-MAY-1999;	99US-0134219P.
PR	25-OCT-1999;	99US-0161406P.	PR	14-MAY-1999;	99US-0134221P.
PR	25-OCT-1999;	99US-0161359P.	PR	18-MAY-1999;	99US-0134370P.
PR	26-OCT-1999;	99US-0161360P.	PR	18-MAY-1999;	99US-0134768P.
PR	26-OCT-1999;	99US-0161361P.	PR	19-MAY-1999;	99US-0134941P.
PR	26-OCT-1999;	99US-0161920P.	PR	20-MAY-1999;	99US-0135124P.
PR	26-OCT-1999;	99US-0161920P.	PR	21-MAY-1999;	99US-0135353P.
PR	28-OCT-1999;	99US-0161922P.	PR	21-MAY-1999;	99US-0135629P.
PR	28-OCT-1999;	99US-0161931P.	PR	25-MAY-1999;	99US-0136021P.
PR	29-OCT-1999;	99US-0162142P.	PR	27-MAY-1999;	99US-0136392P.
PR	29-OCT-1999;	99US-0162142P.	PR	28-MAY-1999;	99US-0136782P.
PR	29-OCT-1999;	99US-0162142P.	PR	01-JUN-1999;	99US-0137222P.
PR	29-OCT-1999;	99US-0162142P.	PR	03-JUN-1999;	99US-0137528P.
PR	29-OCT-1999;	99US-0162142P.	PR	04-JUN-1999;	99US-0137502P.
PR	29-OCT-1999;	99US-0162142P.	PR	07-JUN-1999;	99US-0137724P.
PR	29-OCT-1999;	99US-0162142P.	PR	08-JUN-1999;	99US-0138094P.
PR	29-OCT-1999;	99US-0162142P.	PR	10-JUN-1999;	99US-0138540P.
PR	29-OCT-1999;	99US-0162142P.	PR	10-JUN-1999;	99US-0138847P.
PR	29-OCT-1999;	99US-0162142P.	PR	14-JUN-1999;	99US-0139119P.
PR	29-OCT-1999;	99US-0162142P.	PR	16-JUN-1999;	99US-0139452P.
PR	29-OCT-1999;	99US-0162142P.	PR	16-JUN-1999;	99US-0139453P.
PR	29-OCT-1999;	99US-0162142P.	PR	17-JUN-1999;	99US-0139453P.
PR	29-OCT-1999;	99US-0162142P.	PR	17-JUN-1999;	99US-0139492P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139454P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139455P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139456P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139457P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139458P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139459P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139460P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139461P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139462P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139463P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139750P.
PR	29-OCT-1999;	99US-0162142P.	PR	18-JUN-1999;	99US-0139763P.
PR	29-OCT-1999;	99US-0162142P.	PR	21-JUN-1999;	99US-0139817P.
PR	29-OCT-1999;	99US-0162142P.	PR	22-JUN-1999;	99US-0139899P.
PR	29-OCT-1999;	99US-0162142P.	PR	23-JUN-1999;	99US-0140353P.
PR	29-OCT-1999;	99US-0162142P.	PR	23-JUN-1999;	99US-0140354P.
PR	29-OCT-1999;	99US-0162142P.	PR	24-JUN-1999;	99US-0140695P.
PR	29-OCT-1999;	99US-0162142P.	PR	28-JUN-1999;	99US-0140823P.
PR	29-OCT-1999;	99US-0162142P.	PR	29-JUN-1999;	99US-0140991P.
PR	29-OCT-1999;	99US-0162142P.	PR	30-JUN-1999;	99US-0141287P.
PR	29-OCT-1999;	99US-0162142P.	PR	01-JUL-1999;	99US-0141842P.
PR	29-OCT-1999;	99US-0162142P.	PR	01-JUL-1999;	99US-0142154P.
PR	29-OCT-1999;	99US-0162142P.	PR	02-JUL-1999;	99US-0142055P.
PR	29-OCT-1999;	99US-0162142P.	PR	06-JUL-1999;	99US-0142390P.
PR	29-OCT-1999;	99US-0162142P.	PR	08-JUL-1999;	99US-0142803P.
PR	29-OCT-1999;	99US-0162142P.	PR	09-JUL-1999;	99US-0142920P.
PR	29-OCT-1999;	99US-0162142P.	PR	12-JUL-1999;	99US-0142977P.
PR	29-OCT-1999;	99US-0162142P.	PR	13-JUL-1999;	99US-0143542P.
PR	29-OCT-1999;	99US-0162142P.	PR	14-JUL-1999;	99US-0143624P.
PR	29-OCT-1999;	99US-0162142P.	PR	15-JUL-1999;	99US-0144005P.
PR	29-OCT-1999;	99US-0162142P.	PR	16-JUL-1999;	99US-0144085P.
PR	29-OCT-1999;	99US-0162142P.	PR	16-JUL-1999;	99US-0144085P.
PR	29-OCT-1999;	99US-0162142P.	PR	19-JUL-1999;	99US-0144325P.
PR	29-OCT-1999;	99US-0162142P.	PR	19-JUL-1999;	99US-0144331P.
PR	29-OCT-1999;	99US-0162142P.	PR	19-JUL-1999;	99US-0144332P.
PR	29-OCT-1999;	99US-0162142P.	PR	19-JUL-1999;	99US-0144333P.
PR	29-OCT-1999;	99US-0162142P.	PR	19-JUL-1999;	99US-0144334P.
PR	29-OCT-1999;	99US-0162142P.	PR	19-JUL-1999;	99US-0144335P.
PR	29-OCT-1999;	99US-0162142P.	PR	20-JUL-1999;	99US-0144352P.
PR	29-OCT-1999;	99US-0162142P.	PR	20-JUL-1999;	99US-0144632P.
PR	29-OCT-1999;	99US-0162142P.	PR	20-JUL-1999;	99US-0144684P.
PR	29-OCT-1999;	99US-0162142P.	PR	21-JUL-1999;	99US-0144814P.

Query Match 14.6%; Score 137.5; DB 3; Length 426;
Best Local Similarity 25.0%; Pred. No. 4.6e-05;
Matches 45; Conservative 26; Mismatches 88; Indels 21; Gaps 6;
Qy 1 CPLOGSHALCTCCFPQMPDRRAERQDPVAPQOCACVLCLOPFCHLYMGCTRTGCGCLAP 60
Db 228 CEHGAHLQCHLCQGMPPFRANLQ-----VPLCKGCDPFCGAYWS-SENVTQGVSGP 280
Qy 61 FCELN---LQDKCLDGV---LNNNSYEDILKNYLATRLTKNMLTSL-----VAL 107
Db 281 VCURETERPISERTITRIPETHEMNRHQDITQCIAMKVTVPDVAEWLRLFNREI 340
Qy 108 QRGVFLSDYR-VTGDTVLCVCCGLRFRELTVOYRONIPASELPVAVTSRPPCYWGRNC 166
Db 341 DRSRPLNHAETITASTHVCNDCYKLVGFLLYFRITLPRNHLDPADVAAREDCWTGYAC 400

RESULT 12
AAG20352
ID AAG20352 standard; protein; 453 AA.
XX
AC AAG20352;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22505.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135333P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0132407P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149823P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 12-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158363P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.

PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	22-JUL-1999;	99US-0145145P.
PR	22-JUL-1999;	99US-0145218P.
PR	22-JUL-1999;	99US-0145224P.
PR	22-JUL-1999;	99US-0145276P.
PR	22-JUL-1999;	99US-0145313P.
PR	22-JUL-1999;	99US-0145318P.
PR	22-JUL-1999;	99US-0145351P.
PR	22-JUL-1999;	99US-0145386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159229P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159658P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	22-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 14.6%; Score 137.5; DB 3; Length 426;		
Best Local Similarity 25.0%; Pred. No. 4.6e-05;		
Matches 45; Conservative 26; Mismatches 88; Indels 21; Gaps 6;		
QY	1	CPLOQSHALCTCCFQPMEDRAERQDPRVAPQOCACVCLQFPCHLYMGCTRTGCVGCLAP 60
DB	228	CEHGHAIQCHLCQGMPPFRANLQ-----VPLHCKGCDRPFCCAYWS-SENVTQGVSGP 280
QY	61	FCELN-----LGDKCLDGV-----LNNNSYESDILKNYLATRGLTWKNNMLTESI-----VAL 107
DB	281	VCVRETFPISERTITRIPFTIHENRHEQDITQRCIAHMEKTVDDVVAENLRLFNREI 340
QY	108	QRGVFLSDYR-VGTGTVLCVCCGLRSFRELTYQVRQNI PASLFPVAVTSRDCYWGRC 166
DB	341	DRSRMPLNHAEMITASTHVCDYKLVGFLLYNFRITLPRNHLFPADVAAREDCWYGAC 400
RESULT 11		
AAG37659		
ID	AAG37659	standard; protein; 426 AA.
XX	AC	AAG37659;
XX	XX	
DT	18-OCT-2000	(first entry)
XX	XX	
DE	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 46343.
KW	XX	Protein identification; signal transduction pathway; metabolic pathway;
KW	XX	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	XX	termination sequence.
OS	XX	Arabidopsis thaliana.
XX	XX	
PN	XX	EP1033405-A2.
PD	XX	06-SEP-2000.
XX	XX	
XX	XX	25-FEB-2000; 2000EP-00301439.
XX	XX	
PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.

PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145218P.
PR	26-JUL-1999;	99US-0145276P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145918P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0147038P.
PR	03-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	12-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	17-AUG-1999;	99US-0149426P.
PR	18-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	25-AUG-1999;	99US-0150884P.
PR	26-AUG-1999;	99US-0150865P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152263P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	28-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	08-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158322P.
PR	12-OCT-1999;	99US-0158469P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145218P.
PR	26-JUL-1999;	99US-0145276P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145918P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0147038P.
PR	03-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	12-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	17-AUG-1999;	99US-0149426P.
PR	18-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	25-AUG-1999;	99US-0150884P.
PR	26-AUG-1999;	99US-0150865P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152263P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	28-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	08-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158322P.
PR	12-OCT-1999;	99US-0158469P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145218P.
PR	26-JUL-1999;	99US-0145276P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145918P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	03-AUG-1999;	99US-0147038P.
PR	03-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	12-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	17-AUG-1999;	99US-0149426P.
PR	18-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	25-AUG-1999;	99US-0150884P.
PR	26-AUG-1999;	99US-0150865P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152263P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	28-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	08-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158322P.
PR	12-OCT-1999;	99US-0158469P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.

Query Match

14.6%; Score 137.5; DB 3; Length 350;

Best Local Similarity 25.0%; Pred. No. 3.6e-05;

Matches 45; Conservative 26; Mismatches 88; Indels 21; Gaps 6;

Qy	1	CPLQSHALCTCCFQMPDRRAEREDPRVAPQCCAVLCLOPFCHLYGWCTRTGCGCLAP	60
Db	152	CEHGAHLQCHLCQGMPPFRANLQ-----VPLHCKGCDRPPFCGAYWS-SENVTQGVSGP	204
Qy	61	FCELN-----LGDKCLDGV---LNNNSVESDILKNYLATRGLTWKNMLTESL-----VAL	107
Db	205	VCVRETFRPISERTITRIPITHEMREHQDITQRCIAHMEKTVDPVVAEWLRLFNNREI	264
Qy	108	QRGVFLLSDYR-VTGDVLCYCCGLRSFRLTYQYRONIPASELPVAVTSRDCYWGRC	166
Db	265	DRSRMPLNHAEMITASTHVCNDCYKLVGLVFLYWFRTILPRNHLPAADVAAREDCWGYAC	324

RESULT 9

AAG37660

ID AAG37660 standard; protein; 350 AA.

XX AAG37660;

XX AAG37660;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 46344.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; Genetic mapping; gene expression control; promoter;

KW	termination sequence.
XX	Arabidopsis thaliana.
OS	Arabidopsis thaliana.
PN	EP1033405-A2.
PD	06-SEP-2000.
XX	25-FEB-2000; 2000EP-00301439.
XX	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127482P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI; 2003-395539/38.
DR N-PSDB; ADA52592.
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 1799; 205pp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 128 AA;
Query Match 46.4%; Score 438; DB 6; Length 128;
Best Local Similarity 98.8%; Pred. No. 2.1e-35;
Matches 80; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 86 KNYLATRGLTWKMLTESLVALQGVFLSDYRVTGDTVLCYCGLRSFRELTYQYRNI 145
DB 25 QNYLATRGLTWKMLTESLVALQGVFLSDYRVTGDTVLCYCGLRSFRELTYQYRNI 84
QY 146 PASELPVAVTSRDPYWGRC 166
DB 85 PASELPVAVTSRDPYWGRC 105
RESULT 8
AAG20354
ID AAG20354 standard; protein; 350 AA.
AC AAG20354;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22507.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
DN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 18-JUL-1999; 99US-0144086P.
PR 18-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.

length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 12128; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

Sequence 623 AA;

Query Match 99.6%; Score 939; DB 4; Length 623;
Best Local Similarity 99.4%; Pred. No. 3.5e-84;
Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPLQSHALCTCCFQMPDRRAERQDPRAVQCCAVCLQPFCHLYWGCTRTGCGCLAP 60
DB 435 CPLQSHALCTCCFQMPDRRAERQDPRAVQCCAVCLQPFCHLYWGCTRTGCGCLAP 494
QY 61 FCELNLDGKCLDGVNNSYESDILKNYLATRGLTWKNMLTESLVALQRGVFLLSDYRVT 120
DB 495 FCELNLDGKCLDGVNNSYESDILKNYLATRGLTWKNMLTESLVALQRGVFLLSDYRVT 554
QY 121 GDTVLCYCCGLRSFRELTYQYQNIPASELPVAVTSRPPCYWGRNC 166
DB 555 GDTVLCYCCGLRSFRELTYQYQNIPASELPVAVTSRPPCYWGRNC 600

RESULT 6
ABB97233
ID ABB97233 standard; protein; 623 AA.

AC ABB97233;
XX 27-JUN-2002 (first entry)
XX Novel human protein SEQ ID NO: 501.
XX Human; antianaemic; vulnary; antiinflammatory; immunomodulator;
XX antiinfertility; cerebroprotective; cytotatic; rheumatic; gene therapy;
XX neuroprotective; antiaparkinsonian; protein therapy; EST;
XX expressed sequence tag.
XX Homo sapiens.
XX WO200222660-A2.
XX 21-MAR-2002.
XX 10-SEP-2001; 2001WO-US026015.
XX 11-SEP-2000; 2000US-00659671.
XX

(HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-292408/33.
XX N-PSDB; ABN32419.
XX

An isolated polynucleotide for treating diseases associated with its encoded polypeptide such as cancer and multiple sclerosis.

Example 2; SEQ ID NO 501; 509pp; English.

The present invention provides the protein and coding sequences of 444 novel human proteins. These were isolated from expressed sequences tags (ESTs). They can be used to stimulate cell growth, to regulate haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth e.g. in burn treatment, to regulate the immune system e.g. to treat multiple sclerosis, to regulate activin or inhibin e.g. to treat infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke and cancer, to screen for drugs, to treat inflammatory conditions e.g. rheumatoid arthritis, and to treat nervous system disorders e.g. Parkinson's disease. The present sequence is a protein of the invention

Sequence 623 AA;

Query Match 99.6%; Score 939; DB 5; Length 623;
Best Local Similarity 99.4%; Pred. No. 3.5e-84;
Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPLQSHALCTCCFQMPDRRAERQDPRAVQCCAVCLQPFCHLYWGCTRTGCGCLAP 60
DB 435 CPLQSHALCTCCFQMPDRRAERQDPRAVQCCAVCLQPFCHLYWGCTRTGCGCLAP 494
QY 61 FCELNLDGKCLDGVNNSYESDILKNYLATRGLTWKNMLTESLVALQRGVFLLSDYRVT 120
DB 495 FCELNLDGKCLDGVNNSYESDILKNYLATRGLTWKNMLTESLVALQRGVFLLSDYRVT 554
QY 121 GDTVLCYCCGLRSFRELTYQYQNIPASELPVAVTSRPPCYWGRNC 166
DB 555 GDTVLCYCCGLRSFRELTYQYQNIPASELPVAVTSRPPCYWGRNC 600

RESULT 7
ADA54231
ID ADA54231 standard; protein; 128 AA.

XX ADA54231;
XX 20-NOV-2003 (first entry)
XX Human protein, SEQ ID 1799.
XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
XX Gene Therapy; human; secretory protein; membrane proteins; cancer;
XX inflammatory disease; osteoporosis; neurological disease.
XX Homo sapiens.

XX EP1293569-A2.
XX 19-MAR-2003.
XX 21-MAR-2002; 2002EP-00006586.
XX 14-SEP-2001; 2001JP-00328391.
XX 24-JAN-2002; 2002US-0350435P.
XX (HELI-) HELIX RES INST.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

CC prophase to metaphase during mitosis. It has ubiquitin-protein ligase
CC activity. The Chfr checkpoint was evident in primary human cells, but was
CC inactivated in 4 of 8 human cancer cell lines. In U2OS cells, a mutation
CC was identified that caused a Val to Met amino acid substitution in the
CC highly conserved C-terminal Cys-rich region of the Chfr protein. In the
CC absence of the Chfr checkpoint, cells subjected to mitotic stress
CC condensed their chromosomes despite failing to separate their
CC chromosomes. Chfr may monitor centrosome separation. Inactivation of the
CC Chfr gene (see AAF035) in human cancer is theorized to underlie the
CC increased sensitivity of cancer cells to antimitotic drugs. Polypeptides
CC comprising the present sequence, or sequences comprising at least amino
CC acids 31-103, 303-346 and/or 476-641 of this sequence, are claimed.
CC Claimed methods of determining the tumorigenic potential of a cell
CC comprise examining the cell for the presence of Chfr expression or for
CC the presence of Chfr-mediated ubiquitin-protein ligase activity (in both
CC cases, absence of expression indicating predisposition to tumorigenesis
CC upon exposure to mitotic stress). A diagnostic kit for detecting the
CC tumorigenic potential of cell cells comprises may comprise a ligand that
CC binds to Chfr, such as an antibody or its fragment. Inhibitors of Chfr
CC are identified by monitoring their effect on Chfr expression, and are
CC used to retard the growth of cancer cells
XX
SQ Sequence 564 AA;

Query Match 100.0%; Score 943; DB 4; Length 664;
Best Local Similarity 100.0%; Pred. No. 1.5e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPLQSSHALCTCCFQMPDRRAERQDPRAVQCAVCLQPFCHLYMGCTRTGCGCLAP 60
Db 476 CPLQSSHALCTCCFQMPDRRAERQDPRAVQCAVCLQPFCHLYMGCTRTGCGCLAP 535
QY 61 FCENLGDKCLDGVLNNSYESDILKNYLATRGLTWKMLTESLVALQGVFLLSDYRVT 120
Db 536 FCENLGDKCLDGVLNNSYESDILKNYLATRGLTWKMLTESLVALQGVFLLSDYRVT 595
QY 121 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGRC 166
Db 596 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGRC 641

RESULT 4
AAO08972
ID AAO08972 standard; protein; 269 AA.
XX
AC AAO08972;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 22864.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
FN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI8903.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 22864; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 269 AA;

Query Match 99.7%; Score 940; DB 4; Length 269;
Best Local Similarity 99.4%; Pred. No. 1e-84;
Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPLQSSHALCTCCFQMPDRRAERQDPRAVQCAVCLQPFCHLYMGCTRTGCGCLAP 60
Db 81 CPLQSSHALCTCCFQMPDRRAERQDPRAVQCAVCLQPFCHLYMGCTRTGCGCLAP 140
QY 61 FCENLGDKCLDGVLNNSYESDILKNYLATRGLTWKMLTESLVALQGVFLLSDYRVT 120
Db 141 FCENLGDKCLDGVLNNSYESDILKNYLATRGLTWKMLTESLVALQGVFLLSDYRVT 200
QY 121 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGRC 166
Db 201 GDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRDPDCYWGRC 246

RESULT 5
AAB93182
ID AAB93182 standard; protein; 623 AA.
XX
AC AAB93182;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12128.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
FN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Iehii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-

CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 652 AA;

Query Match 100.0%; Score 943; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 1.5e-84; Indels 0; Gaps 0;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPLQSHALCTCCFQMPDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 60
Db 464 CPLQSHALCTCCFQMPDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 523
QY 61 FCELNLDKCLDGLVNNNSYESDILKNYLATRGLTWKNNMLTESLVALQRGVFLLSDYRVT 120
Db 524 FCELNLDKCLDGLVNNNSYESDILKNYLATRGLTWKNNMLTESLVALQRGVFLLSDYRVT 583
QY 121 GDTVLCVCCGLRSFRELTYQYRQNPASELPVAVTSRPPDCYWGRC 166
Db 584 GDTVLCVCCGLRSFRELTYQYRQNPASELPVAVTSRPPDCYWGRC 629

RESULT 2
AAB83843
ID AAB83843 standard; protein; 664 AA.
XX AAB83843;
AC AAB83843;
XX
XX 22-AUG-2001 (first entry)
DT Amino acid sequence of a human ring finger protein designated FHAR1.
DE FHAR1; RING finger protein; cancer; vaccine.
XX
XX Homo sapiens.
OS
XX
XX WO200142430-A1.
PN
XX
XX 14-JUN-2001.
PD
XX
XX 07-DEC-2000; 2000WO-US033094.
PF
XX
XX 08-DEC-1999; 99US-00456876.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA
XX
XX Zhou B, Zhu Y, Chaturvedi P, Hurler MR, Li X;
PI
XX
XX WPI; 2001-381663/40.
DR
XX
XX N-PSDB; AAF89709.
DR
XX
XX New FHAR1 polypeptide, a member of the RING finger protein family for
PT diagnosing and treating cancer, and for use in anti-cancer vaccines.
PT
XX
XX Claim 1; Page 19; 28pp; English.
PS
XX
XX The present sequence represents a FHAR1 polypeptide, which is a member of
CC the RING finger protein family. FHAR1 is useful in the treatment of
CC cancer, and as a vaccine for inducing an immunological response in a
CC mammal. FHAR1 polynucleotides may also be used as a diagnostic reagent
CC through detection of mutations in the associated gene, and for chromosome

CC localization studies, and tissue expression studies. FHAR1 antibodies are
CC useful to isolate and to identify clones expressing the polypeptides, or
CC to purify the polypeptides by affinity chromatography and to treat cancer
XX
SQ Sequence 664 AA;
Query Match 100.0%; Score 943; DB 4; Length 664;
Best Local Similarity 100.0%; Pred. No. 1.5e-84; Indels 0; Gaps 0;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPLQSHALCTCCFQMPDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 60
Db 476 CPLQSHALCTCCFQMPDRAEREDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAP 535
QY 61 FCELNLDKCLDGLVNNNSYESDILKNYLATRGLTWKNNMLTESLVALQRGVFLLSDYRVT 120
Db 536 FCELNLDKCLDGLVNNNSYESDILKNYLATRGLTWKNNMLTESLVALQRGVFLLSDYRVT 595
QY 121 GDTVLCVCCGLRSFRELTYQYRQNPASELPVAVTSRPPDCYWGRC 166
Db 596 GDTVLCVCCGLRSFRELTYQYRQNPASELPVAVTSRPPDCYWGRC 641

RESULT 3
AAB20219
ID AAB20219 standard; protein; 664 AA.
XX AAB20219;
AC AAB20219;
XX
XX 14-MAY-2001 (first entry)
DT Human Chfr (checkpoint with FHA and ring finger) protein.
DE Checkpoint with forkhead associated domain and ring finger; Chfr; human;
KW mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening;
KW ubiquitin-protein ligase.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Domain 31..103
FT Domain /label= Forkhead-associated_domain
FT Domain 303..346
FT Domain /label= Ring_finger_domain
FT Region 476..641
FT Misc-difference 580 /note= "cysteine-rich region"
FT /note= "Met in U2OS cells"

XX WO200109150-A2.
PN
XX
XX 08-FEB-2001.
PD
XX
XX 14-JUN-2000; 2000WO-US016391.
PF
XX
XX 29-JUL-1999; 99US-0146194P.
PR
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA
XX
XX Halazonetis T, Scolnick D;
PI
XX
XX WPI; 2001-182927/18.
DR
XX
XX N-PSDB; AAF30352.
DR
XX
XX Novel nucleic acid sequence of mitotic checkpoint gene encoding a
PT checkpoint with forkhead-associated domain and ring finger protein, for
PT diagnosing tumorigenic cells and in screening for anticancer drugs.
PT
XX
XX Claim 8(a); Fig 4A-C; 85pp; English.
PS
XX
XX The present sequence is that of human mitotic checkpoint protein Chfr.
CC having a forkhead associated domain (FHA) and a ring finger domain. The
CC protein is required for regulation of the transition of cells from

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:36:57 ; Search time 34.0063 Seconds
(without alignments)
1379.240 Million cell updates/sec

Title: US-10-048-046-2_COPY_476_641

Perfect score: 943
Sequence: 1 CPGQSHALCTCCQPMFDR.....ASELPVAVTSRPPCYWGRNC 166

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_25Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	943	100.0	652	4	AAB93168 Human pro
2	943	100.0	664	4	AAB83843 Amino aci
3	943	100.0	664	4	AAB20219 Human Chf
4	940	99.7	269	4	AAB08972 Human pol
5	939	99.6	623	4	AAB93182 Human pro
6	939	99.6	623	5	AAB97233 Novel hum
7	438	46.4	128	6	ADA54231 Human pro
8	137.5	14.6	350	3	AAG20354 Arabidops
9	137.5	14.6	350	3	AAG37660 Arabidops
10	137.5	14.6	426	3	AAG20353 Arabidops
11	137.5	14.6	426	3	AAG37659 Arabidops
12	137.5	14.6	453	3	AAG20352 Arabidops
13	137.5	14.6	453	3	AAG37658 Arabidops
14	92.5	9.8	227	3	AAB34759 Human sec
15	85.5	9.1	244	4	AAB54425 Propionib
16	85.5	9.1	244	6	ABM52944 Propionib
17	85	9.0	229	4	AAM40223 Human pol
18	85	9.0	229	5	ABG34856 Human can
19	85	9.0	231	4	AAM42009 Human pol
20	85	9.0	504	4	AAE02617 Human hae
21	84	8.9	232	6	ABG73844 Mutant hu
22	82.5	8.7	1023	4	ABH11436 D. discoi
23	81	8.6	551	5	ABB04672 Desulfito
24	80	8.5	192	6	ABR55481 Amino aci
25	80	8.5	302	3	AAG47210 Arabidops

26 80 8.5 324 3 AAG47209 Arabidops
27 80 8.5 530 4 ABB71303 Arabidops
28 80 8.5 725 4 AEG18306 Novel hum
29 80 8.5 725 4 AEG28511 Novel hum
30 79.5 8.4 577 5 ABG96264 Human imm
31 79 8.4 317 7 ADE07954 Novel pro
32 79 8.4 1561 5 ABB97437 Novel hum
33 78.5 8.3 493 4 AAE02618 Human hae
34 77 8.2 232 6 ABG73838 Mutant hu
35 77 8.2 303 6 ABU92033 Human pro
36 76.5 8.1 418 4 AAB93353 Human pro
37 76.5 8.1 1700 4 ABB64608 Arabidops
38 76.5 8.1 2931 4 ABB68229 Arabidops
39 76 8.1 185 6 ADA48586 Rice prot
40 76 8.1 185 6 ADA47952 Rice prot
41 76 8.1 567 3 AAY54838 Human dia
42 76 8.1 567 4 AAE01864 Human dia
43 76 8.1 567 6 AAE36090 Human dia
44 76 8.1 567 6 AAE36089 Human dia
45 76 8.1 567 7 ADC77639 Human 994

ALIGNMENTS

RESULT 1
AAB93168
ID AAB93168 standard; protein; 652 AA.
XX
AC AAB93168;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12100.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 95JP-00248036.
PR 27-AUG-1999; 95JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
(HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
WPI: 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the polynucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end

Search completed: May 7, 2004, 14:49:04
Job time : 8.36536 secs

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RN Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AY087766; AAM65302.1; Znf_ring.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 426 AA; 47807 MW; 48974F27881BB833 CRC64;

Query Match 42.7%; Score 113.5; DB 10; Length 426;
Best Local Similarity 46.5%; Pred. No. 5.3e-07;
Matches 20; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 CIIICQLLHDCVSLQPCMHHTFCAACYSGWMERSLICTPCRCVP 44
DB 365 CAICQEKHMTILLR-CKHMFCDVCSEWFERETCPLCALV 406

RESULT 13
Q8L610 PRELIMINARY; PRT; 426 AA.
ID Q8L610;
AC Q8L610;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN AT5G01960.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AY099711; AAM20562.1; --
DR EMBL; BT000282; AAN15601.1; --
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.

DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 426 AA; 47791 MW; B245F521B9F663D1 CRC64;

Query Match 42.7%; Score 113.5; DB 10; Length 426;
Best Local Similarity 46.5%; Pred. No. 5.3e-07;
Matches 20; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 CIIICQLLHDCVSLQPCMHHTFCAACYSGWMERSLICTPCRCVP 44
DB 365 CAICQEKHMTILLR-CKHMFCDVCSEWFERETCPLCALV 406

RESULT 14
Q37928 PRELIMINARY; PRT; 498 AA.
ID Q37928;
AC Q37928;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Immediate early protein.
OS Feline herpesvirus (Feline herpesvirus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-27;
RA Sussman M.D., Maes R.K.;
RT "Nucleotide sequence and characterization of feline herpesvirus 1
RT homologs for ICP0, glycoprotein gL and uracil DNA glycosylase.";
RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AF022391; AAB80763.1; --
DR HSPSP; P28990; ICHC.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 498 AA; 55012 MW; 68C58DED3B9C59C8 CRC64;

Query Match 41.4%; Score 110; DB 12; Length 498;
Best Local Similarity 39.5%; Pred. No. 1.8e-06;
Matches 17; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

QY 2 CIIICQLLHDCVSLQPCMHHTFCAACYSGWMERSLICTPCRCVP 44
DB 8 CPICLDMNDLTFTMPLCHKFCYSLRWVGLNNKCLCKTSV 50

RESULT 15
O13628 PRELIMINARY; PRT; 304 AA.
ID O13628;
AC O13628;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-MAR-2003 (TrEMBLrel. 05, Last sequence update)
DE PAS4 protein.
GN P1036.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972 h-;
RX MEDLINE=20089027; PubMed=10620777;
RA Machida M., Yamazaki S., Kunihiro S., Tanaka T., Jinno K.,
RA Haikawa Y., Yamazaki J., Yamamoto S., Sekine M., Oguchi A., Nagai Y.,


```
RT      60,770 full-length cDNAs." ;  
RL      Nature 420:563-573 (2002).  
RN      [2]  
RP      SEQUENCE FROM N.A.  
RC      STRAIN=FVB/N-3; TISSUE=Breast tumor;  
RA      Strausberg R.;  
DR      Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.  
DR      EMBL; AX089091; BAC40745.1; -  
DR      EMBL; BC043266; AAH43266.1; -  
DR      MGD; MGI:1916837; 231003SNI.Srik.  
DR      InterPro; IPR001841; Znf_ring.  
DR      Pfam; PF000097; zf-C3HC4; 1.  
DR      SMART; SM00184; RING; 1.  
DR      PROSITE; PS00518; ZF_RING_1; 1.  
DR      PROSITE; PSS0089; ZF_RING_2; 1.  
SQ      SEQUENCE 241 AA; 28046 MW; 1F4645C6E7A3DEA5 CRC64;  
  
Query Match          43.2%; Score 115; DB 11; Length 241;  
Best Local Similarity 41.5%; Pred.No. 2e-07;  
Matches 17; Conservative  
  
Qy      1 TCICODLLHDCVSLPCMHFTCAACYSGWMERSSLCPTCR 41  
         |||::|||::|||::|||::|||::|||::|||:  
Db      16 TCRLCGSLYDATTVTTECLHTFCRSLVKYLENNTCPTCR 56  
         |||::|||::|||::|||::|||::|||:  
  
RESULT 11  
OL5262 PRELIMINARY; PRT; 247 AA.  
AC AC OL5262;  
DT DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)  
DD DD 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE DE Ring finger protein.  
GN GN RNF3A.  
OS OS Homo sapiens (Human).  
OC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX OX NCBI_TaxID=9606;  
RN RN [1]  
RP RP SEQUENCE FROM N.A.  
RA RA Abdul-Rauf M., Dyer M.J.;  
RT RT "Interactions of the BCLTA protein with novel ring finger pr  
RL RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.  
RC RC -I- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR DR EMBL; AJ001019; CAA04477.1; -  
DR DR InterPro; IPR001841; Znf_ring.  
DR DR Pfam; PF000097; zf-C3HC4; 1.  
DR DR SMART; SM00184; RING; 1.  
DR DR PROSITE; PS00518; ZF_RING_1; 1.  
DR DR PROSITE; PSS0089; ZF_RING_2; 1.  
KW KW Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 247 AA; 28748 MW; AFA425732644EF8A CRC64;  
  
Query Match          43.2%; Score 115; DB 4; Length 247;  
Best Local Similarity 41.5%; Pred.No. 2.le-07;  
Matches 17; Conservative  
  
Qy      1 TCICODLLHDCVSLPCMHFTCAACYSGWMERSSLCPTCR 41  
         |||::|||::|||::|||::|||::|||:  
Db      21 TCRLCGSLYDATTVTTECLHTFCRSLVKYLENNTCPTCR 61  
         |||::|||::|||::|||::|||::|||:  
  
RESULT 12  
OQLAKO PRELIMINARY; PRT; 426 AA.  
ID ID OQLAKO;  
AC AC OQLAKO;  
DT DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE DE Hypothetical protein.  
OS OS Arabidopsis thaliana (Mouse-ear cress).  
OC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheo
```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20388685; PubMed=10935642;
RA Scolnick D.M., Halazonetis T.D.;
RT "Chfr defines a mitotic stress checkpoint that delays entry into
RT metaphase";
RL Nature 406:430-435(2000).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AF170724; AAF91084.1; -.
DR PDB; 1LGP; 07-AUG-02.
DR PDB; 1LGO; 07-AUG-02.
DR Genew; HGNC:20455; CHFR.
DR GO; GO:0007093; P:mitotic checkpoint; TAS.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
DR Metal-binding; Zinc; Zinc-finger.
KW SEQUENCE 664 AA; 73386 MW; 141A1E7FEFAE36A2 CRC64;
SQ
Query Match 100.0%; Score 266; DB 4; Length 664;
Best Local Similarity 100.0%; Pred. No. 5.1e-27;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCIIQDQLHDCVSLQPCMTTFCACYSGWMSRLCPTCRCPV 44
DB 303 TCIIQDQLHDCVSLQPCMTTFCACYSGWMSRLCPTCRCPV 346
RESULT 6
Q810L3 PRELIMINARY; PRT; 664 AA.
ID Q810L3;
AC Q810L3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RIKEN CDNA 573048420 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Limb;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049792; AAF49792.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 2.
DR PROSITE; PS50006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 664 AA; 73971 MW; D651BE3B463DEB56 CRC64;
Query Match 100.0%; Score 266; DB 11; Length 664;
Best Local Similarity 100.0%; Pred. No. 5.1e-27;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCIIQDQLHDCVSLQPCMTTFCACYSGWMSRLCPTCRCPV 44
DB 302 TCIIQDQLHDCVSLQPCMTTFCACYSGWMSRLCPTCRCPV 345

RESULT 7
Q9SX88 PRELIMINARY; PRT; 473 AA.
ID Q9SX88;
AC Q9SX88;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F16N3.15.
GN F16N3.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J., Liu A., Li J., Kremenetskaia I., Gonzalez A., Altafi H.,
RA Araujo R., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,
RA Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C., Rowley D.,
RA Shinn P., Walker M., Davis R.W., Ecker J.R., Federspiel N.A.,
RA Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F16N3 sequence.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AC007519; AAD46042.1; -.
DR PIR; C96516; C96516.
DR InterPro; IPR001841; Znf ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
DR Metal-binding; Zinc; Zinc-finger.
KW SEQUENCE 473 AA; 53874 MW; D17217493A81C991 CRC64;
SQ
Query Match 50.8%; Score 135; DB 10; Length 473;
Best Local Similarity 46.9%; Pred. No. 8.3e-10;
Matches 23; Conservative 5; Mismatches 15; Indels 6; Gaps 1;
QY 2 CIIICDQLHDCVSLQPCMTTFCACYSGWMSRLCPTCRCPV 44
DB 145 CSICLNIWHDVVVTAAPCLHNFNGCFSEMMRSEKHKHVLCPQCRTV 193
RESULT 8
Q8NJ06 PRELIMINARY; PRT; 829 AA.
ID Q8NJ06;
AC Q8NJ06;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN SP3.020.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AL807371; CAD37003.1; -.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.

DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS0006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 652 AA; 72058 MW; 1A50773CEA679F07 CRC64;

Query Match 100.0%; Score 266; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 5.1e-27; Indels 0; Gaps 0;
Matches 44; Conservative 0; Mismatches 0;

Qy 1 TCIIQDHLHDCVSLQPCMHFTCAACYSGWMERSLCTCRCPV 44
Db 291 TCIIQDHLHDCVSLQPCMHFTCAACYSGWMERSLCTCRCPV 334

RESULT 4
Q8BJZ9 PRELIMINARY; PRT; 663 AA.
ID Q8BJZ9 PRELIMINARY; PRT; 663 AA.
AC Q8BJZ9; 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to cell cycle checkpoint protein CHFR.
GN 5730484M20RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK077629; BAC36912.1; -.
DR MGI; MGI:2444898; 5730484M20RIK.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; ZF-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 2.
DR PROSITE; PS0006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 663 AA; 73855 MW; 42BD653DCB4ADF45 CRC64;

Query Match 100.0%; Score 266; DB 11; Length 663;
Best Local Similarity 100.0%; Pred. No. 5.1e-27; Indels 0; Gaps 0;
Matches 44; Conservative 0; Mismatches 0;

Qy 1 TCIIQDHLHDCVSLQPCMHFTCAACYSGWMERSLCTCRCPV 44
Db 302 TCIIQDHLHDCVSLQPCMHFTCAACYSGWMERSLCTCRCPV 345

RESULT 5
Q9NRT4 PRELIMINARY; PRT; 664 AA.
ID Q9NRT4 PRELIMINARY; PRT; 664 AA.
AC Q9NRT4; 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 25, Last annotation update)
DE Cell cycle checkpoint protein CHFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Qy 1 TCIIQDHLHDCVSLQPCMHFTCAACYSGWMERSLCTCRCPV 44
Db 262 TCIIQDHLHDCVSLQPCMHFTCAACYSGWMERSLCTCRCPV 305

RESULT 2
Q96SL3 PRELIMINARY; PRT; 652 AA.
ID Q96SL3 PRELIMINARY; PRT; 652 AA.
AC Q96SL3; 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ14781.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK027687; BAB55297.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; ZF-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS0006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 652 AA; 72030 MW; 572F2CE6D1743D80 CRC64;

Query Match 100.0%; Score 266; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 5.1e-27; Indels 0; Gaps 0;
Matches 44; Conservative 0; Mismatches 0;

Qy 1 TCIIQDHLHDCVSLQPCMHFTCAACYSGWMERSLCTCRCPV 44
Db 291 TCIIQDHLHDCVSLQPCMHFTCAACYSGWMERSLCTCRCPV 334

RESULT 3
Q96EP1 PRELIMINARY; PRT; 652 AA.
ID Q96EP1 PRELIMINARY; PRT; 652 AA.
AC Q96EP1; 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; BC012072; AAH12072.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; ZF-C3HC4; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:40:47 ; Search time 6.36536 Seconds
(without alignments)
2180.991 Million cell updates/sec

Title: US-10-048-046-2_COPY_303_346

Perfect score: 266
Sequence: 1 TCICQDLLHDCVSLQPCMH.....ACYSGMWERSLCTCRCPV 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25;*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	266	100.0	623	4	Q9NVD5	Q9nvd5 homo sapien
2	266	100.0	652	4	Q96SL3	Q96sl3 homo sapien
3	266	100.0	652	4	Q96EP1	Q96ep1 homo sapien
4	266	100.0	663	11	Q8BJZ9	Q8bjz9 mus musculu
5	266	100.0	664	4	Q9NRT4	Q9nrt4 homo sapien
6	266	100.0	664	11	Q810L3	Q810l3 mus musculu
7	135	50.8	473	10	Q9SX88	Q9sx88 arabidopsis
8	116	44.4	829	3	Q8NJ06	Q8nj06 neurospora
9	115	43.2	178	11	Q8BTZ0	Q8btz0 mus musculu
10	115	43.2	241	11	Q8BTQ0	Q8btq0 mus musculu
11	115	43.2	247	4	O15262	O15262 homo sapien
12	113.5	42.7	426	10	Q8LAK0	Q8lak0 arabidopsis
13	113.5	42.7	426	10	Q8L610	Q8l610 arabidopsis
14	110	41.4	498	12	O37928	O37928 feline herp
15	107.5	40.4	304	3	O13628	O13628 schizosacch
16	107.5	40.4	306	3	Q9UUF0	Q9uuf0 schizosacch

17	107	40.2	551	3	Q08109	Q08109 saccharomyc
18	105.5	39.7	515	10	Q9SIU7	Q9siu7 arabidopsis
19	104	39.1	610	5	Q20798	Q20798 caenorhabdi
20	104	39.1	618	13	Q80318	Q80318 brachydanio
21	104	39.1	639	5	Q21641	Q21641 caenorhabdi
22	103.5	38.9	540	13	Q7ZX20	Q7zx20 xenopus lae
23	103	38.7	298	4	Q8TEH6	Q8teh6 homo sapien
24	103	38.7	575	11	Q80T88	Q80t88 mus musculu
25	103	38.7	579	4	Q8GJL5	Q8gj15 homo sapien
26	103	38.7	612	11	Q8GB5	Q8gb5 mus musculu
27	103	38.7	612	11	Q9DBY1	Q9dbyl mus musculu
28	103	38.7	616	4	Q8N3K3	Q8n3k3 homo sapien
29	103	38.7	616	4	Q96PK3	Q96pk3 homo sapien
30	103	38.7	617	4	Q8N6E8	Q8n6e8 homo sapien
31	103	38.7	617	4	Q86TW6	Q86tm6 homo sapien
32	102	38.3	328	5	Q9VJW5	Q9vjw5 drosophila
33	102	38.3	351	5	Q9NKF3	Q9nkf3 drosophila
34	101.5	38.2	161	10	Q9FN58	Q9fn58 arabidopsis
35	101.5	38.2	237	4	Q96A37	Q96a37 homo sapien
36	101.5	38.2	459	3	Q7ZAI7	Q7zai7 yarrowia li
37	101	38.0	496	11	Q9D6L2	Q9d6l2 mus musculu
38	101	38.0	587	5	Q96182	Q96182 plasmodium
39	100.5	37.8	132	10	Q9M0R7	Q9m0r7 arabidopsis
40	100.5	37.8	486	3	O60106	O60106 schizosacch
41	100.5	37.8	496	10	Q9SS55	Q9ss55 arabidopsis
42	100.5	37.8	652	4	Q86Y15	Q86y15 homo sapien
43	100.5	37.8	673	12	Q67469	Q67469 feldmannia
44	100.5	37.8	1208	4	O75162	O75162 homo sapien
45	100	37.6	168	12	Q8QLD7	Q8qld7 mamestra co

ALIGNMENTS

RESULT 1

ID	Q9NVD5	PRELIMINARY;	PRT;	623 AA.
AC	Q9NVD5;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Hypothetical protein FLJ10796.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,			
RA	Matsumawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,			
RA	Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;			
RT	"NEDO human cDNA sequencing project."			
RL	Submitted (PEB-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.			
DR	EMBL: AK001658; BAA91817.1; -			
DR	InterPro; IPR000253; FHA.			
DR	InterPro; IPR008984; SMAD_FHA.			
DR	InterPro; IPR001841; Znf_fing.			
DR	Pfam; PF00498; FHA; 1.			
DR	Pfam; PF00097; zf-C3HC4; 1.			
DR	SMART; SM00240; FHA; 1.			
DR	SMART; SM00184; RING; 1.			
DR	PROSITE; PS00066; FHA_DOMAIN; 1.			
DR	PROSITE; PS00518; ZF_RING_1; 1.			
DR	PROSITE; PS00089; ZF_RING_2; 1.			
KW	Hypothetical protein; Metal-binding; Zinc; Zinc-finger.			
SQ	SEQUENCE 623 AA; 69204 MW; 45186D33DAE52711 CRC64;			

Query Match 100.0%; Score 266; DB 4; Length 623;
Best Local Similarity 100.0%; Pred. No. 4.9e-27;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 6; Mismatches 21; Indels 2; Gaps 1;
 ID TRM3_MOUSE STANDARD; PRT; 744 AA.
 AC Q9R1R2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Tripartite motif protein 3 (RING finger protein 22) (RING finger
 protein HACL1).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Yanai K., Shimamoto Y., Hirota K., Fukamizu A.;
 RT "Cloning of a new co-activator with ring finger motif.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21231161; PubMed=11331580;
 RA Raymond A., Meroni G., Pantozi A., Merla G., Cairo S., Luzi L.,
 RA Riganelli D., Zanaria E., Messori S., Calmarca S., Guffanti A.,
 RA Minucci S., Pelicci P.G., Ballabio A.;
 RT "The tripartite motif family identifies cell compartments.";
 RL EMO J. 20:2140-2151(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RN Nature 409:685-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton B., Kettman M., Madao A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- SUBUNIT: Associates with myosin V and alpha-actinin-4 (By
 similarity).
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -1- SIMILARITY: Contains 1 B box-type zinc finger.
 CC -1- SIMILARITY: Contains 1 filament repeat.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB030312; BAA83343.1; -
 CC EMBL; AK220019; AAG53473.1; -
 CC EMBL; AK019165; BAB31580.1; -
 CC EMBL; BC034263; AAR34263.1; -
 CC HSSP; P29590; 1BOR.
 CC MGD; MGI:1860040; Trim3.
 CC GO; GO:0005737; C:cytoplasm; IDA.
 CC GO; GO:0005515; F:protein binding; IPI.
 CC InterPro; IPR003649; Bbox_C.
 CC InterPro; IPR001298; Filamin.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR001258; NHL.
 CC InterPro; IPR000315; Znf_Bbox.
 CC InterPro; IPR001841; Znf_Ring.
 CC Pfam; PF00630; Filamin; 1.
 CC Pfam; PF01436; NHL; 6.
 CC Pfam; PF00643; zf-B_box; 1.
 CC Pfam; PF00037; zf-C3HC4; 1.
 CC SMART; SM00502; BBC; 1.
 CC SMART; SM00336; BBOX; 1.
 CC SMART; SM00557; IG_FLMN; 1.
 CC SMART; SM00184; RING; 1.
 CC PROSITE; PSS0194; FILAMIN_REPEAT; 1.
 CC PROSITE; PSS0119; ZF_BBOX; 1.
 CC PROSITE; PSS0518; ZF_RING_1; 1.
 CC PROSITE; PSS0089; ZF_RING_2; 1.
 CC Zinc-finger; Coiled Coil.
 KW ZN FING; 22 63 RING-TYPE.
 FT ZN FING 110 151 B_BOX-TYPE.
 FT DOMAIN 153 224 COILED COIL (POTENTIAL).
 FT REPEAT 317 418 FILAMIN.
 SQ SEQUENCE 744 AA; 80774 MW; D9AEF4FA264BA168 CRC64;
 Query Match 36.1%; Score 96; DB 1; Length 744;
 Best Local Similarity 41.9%; Pred. No. 0.00078;
 Matches 18; Conservative 5; Mismatches 16; Indels 4; Gaps 2;
 QY 2 CIIQCDDLLHDCVSLQPCMHFTFCACVSGWMMERSL----CPTCR 41
 DB 22 CSICLD-RYRCPKVLPCILHTFCERCLQNYIPQSLTSLSCPVC 63

Search completed: May 7, 2004, 14:46:33
 Job time : 2.4868 secs

ID AC Q9YK7; Q9E011; STANDARD; PRT; 154 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RING finger protein 11 (Sidi669) (Nedd4 WW domain-binding protein 2).
GN RNF11 OR N4WBP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20135600; PubMed=10673045;
RA Seki N., Hattori A., Hayashi A., Kozuma S., Sasaki M., Suzuki Y.,
RA Sugano S., Muramatsu M., Saigo T.,
RT "Cloning and expression profile of mouse and human genes, Rnf11/RNF11,
RT encoding a novel RING-H2 finger protein";
RL Biochim. Biophys. Acta 1489:421-427(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20498735; PubMed=11042109;
RA Jolliffe C.N., Harvey K.F., Haines B.P., Parasivam G., Kumar S.;
RT "Identification of multiple proteins expressed in murine embryos as
RT binding partners for the WW domains of the ubiquitin-protein ligase
RT Nedd4";
RL Biochem. J. 351:557-565(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gnaratine P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny K.C., Harte S., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S., Krywinski W.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra W.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBUNIT: Interacts with NEDD4.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; AB024427; BAA94682.1; -
CC EMBL; AF220206; AAG44245.1; ALT_INIT.
CC EMBL; BC010299; AAH10299.1; -
CC MGD; MGI:135759; Rnf11.
CC GO; GO:000151; C:ubiquitin ligase complex; IPI.
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IPI.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00097; zf-C3HC4; 1.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS50089; ZF_RING_1; FALSE_NEG.
CC PROSITE; PS50089; ZF_RING_2; 1.

KW Zinc-finger. RING-TYPE.
FT ZN_FING 99 140
SQ SEQUENCE 154 AA; 17458 MW; EF192AB0C2D4BF87 CRC64;
Query Match 36.5%; Score 97; DB 1; Length 154;
Best Local Similarity 40.0%; Pred. No. 0.00015;
Matches 18; Conservative 7; Mismatches 18; Indels 2; Gaps 2;
QY 2 CIIIC-QDLH-DCVSLQFCMHTFCAACYSGWMERSLCTPCPCPV 44
DB 99 CVICWMDPVYGDPIRFLPCVHYHLDICDDWLKRSFTCPSCMEPV 143
RESULT 14
ICP0_HSV2H
ID ICP0_HSV2H STANDARD; PRT; 825 AA.
AC P28284;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-acting transcriptional protein ICP0 (VMW118 protein).
GN RL2.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2";
RL J. Gen. Virol. 72:3057-3075(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ICP0 PROTEIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; D10471; BAA23427.1; -
CC EMBL; Z86099; CAB06760.1; -
CC PIR; JQ1501; EDBEXD.
CC HSP; P28990; 1CHC.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00097; zf-C3HC4; 1.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS50089; ZF_RING_2; 1.
KW Transcription regulation; Trans-acting factor; Activator; Zinc-finger;
KW DNA-binding.
FT DOMAIN 120 123 POLY-GLY.
FT ZN_FING 126 167 RING-TYPE.
FT DOMAIN 266 271 POLY-GLY.
FT DOMAIN 292 295 POLY-SER.
FT DOMAIN 342 345 POLY-ALA.
FT DOMAIN 386 389 POLY-SER.
FT DOMAIN 395 400 POLY-GLY.
FT DOMAIN 425 428 POLY-ALA.
FT DOMAIN 590 627 POLY-SER.
SQ SEQUENCE 825 AA; 81986 MW; 5CEB15858553A274 CRC64;
Query Match 36.5%; Score 97; DB 1; Length 825;
Best Local Similarity 35.6%; Pred. No. 0.00066;

[illegible]

FT SQ SEQUENCE 744 AA; 80836 MW; 3338151B7368B37E CRC64; /FTID=VSP_005759.
Query Match 37.2%; Score 99; DB 1; Length 744;
Best Local Similarity 41.9%; Pred. No. 0.00036;
Matches 18; Conservative 5; Mismatches 16; Indels 2; Gaps 2;
OY 2 CIIQDLHDCVSLQPCVMTFCAACYSQWMSRL---CPTCR 41
DB 22 CSICLD-RYQCPKVLPLHTFCERCLQNVIPAQSILTSFCVCR 63
RESULT 10
CBLC HUMAN STANDARD; PRT; 474 AA.
AC Q9ULV8; O9Y5Z2; O9Y5Z3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transduction protein CBL-C (SH3-binding protein CBL-C) (CBL-3).
GN CBLC OR CBL3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM LONG).
RP MEDLINE=20035821; PubMed=10571044;
RX Kim M., Tezuka T., Suzuki Y., Sugano S., Hirai M., Yamamoto T.;
RA "Molecular cloning and characterization of a novel cbl-family gene,
RT cbl-c".
RL Gene 239:145-154 (1999).
[2]
SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RP TISSUE=Pancreatic adenocarcinoma;
RX MEDLINE=92829203; PubMed=10362357;
RA Keane M.M., Ettenberg S.A., Nau M.M., Banerjee P., Cuello M.,
RA Penninger J., Lipkowitz S.;
RT "Cbl-3: a new mammalian cbl family protein.";
RL Oncogene 18:3365-3375 (1999).
CC -1- FUNCTION: Regulator of EGFR mediated signal transduction.
CC -1- SUBUNIT: Interacts with a restricted range of SH3 domain proteins.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Names=Long;
CC IsoId=Q9ULV8-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q9ULV8-2; Sequence=VSP_005732;
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- DOMAIN: The N-terminus is composed of the phosphotyrosine binding
(PTB) domain, a short linker region and the RING-type zinc finger.
CC The PTB domain, which is also called TKB (tyrosine kinase binding)
CC domain, is composed of three different subdomains: a four-helix
CC bundle (4H), a calcium-binding EF hand and a divergent SH2 domain.
CC -1- DOMAIN: The RING-type zinc finger domain mediates binding to an E2
CC ubiquitin-conjugating enzyme (By similarity).
CC -1- PTM: phosphorylated on tyrosines by EGFR.
CC -1- MISCELLANEOUS: This protein has one functional calcium-binding
CC site (By similarity).
CC -1- SIMILARITY: Contains 2 EF-hand-like calcium-binding domains.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: Contains 1 SH2 domain.
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobogen.fr/services/chromancer/Genes/CBLCID194.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

CC EMBL; AB028645; BAA8298.1; -;
DR EMBL; AF117646; AAD34341.1; -;
DR EMBL; AF117647; AAD34342.1; -;
DR HSSP; P22681; 1B47.
DR Genew; HGNC:15961; CBLC.
DR GO; GO:0005515; F:protein binding; TAS.
DR GO; GO:0008270; F:zinc ion binding; TAS.
DR GO; GO:0007175; P:negative regulation of EGF receptor activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR InterPro; IPR003153; Cbl_N.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF02761; Cbl_N2; 1.
DR Pfam; PF02762; Cbl_N3; 1.
DR Pfam; PF02822; Cbl_N; 1.
DR Pfam; PF02822; Cbl_N; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Nuclear protein; Zing-finger; Phosphorylation; Alternative splicing;
KW SH2 domain; Calcium-binding. PTB.
FT DOMAIN 8 327
FT DOMAIN 8 327
FT DOMAIN 8 327
FT DOMAIN 182 190 EF-HAND LIKE 1.
FT DOMAIN 182 190 EF-HAND LIKE 2.
FT CA_BIND 199 210 SH2.
FT DOMAIN 237 311 LINKER.
FT DOMAIN 312 350 RING-TYPE.
FT ZN_FING 351 390 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 89 92 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT BINDING 264 264 PHOSPHOTYROSINE (BY SIMILARITY).
FT BINDING 261 306 Missing (in isoform Short).
FT VARSPLIT 261 306 /FTID=VSP_005732.
FT CONFLICT 234 234 T -> N (IN REF. 2).
FT SEQUENCE 474 AA; 52468 MW; 91013DDF12828242 CRC64;
Query Match 36.8%; Score 98; DB 1; Length 474;
Best Local Similarity 38.6%; Pred. No. 0.00031;
Matches 17; Conservative 6; Mismatches 19; Indels 2; Gaps 2;
OY 2 CIIQDLHDCVSLQPCVMTFCAACYSQWMSRL---CPTCR 44
DB 351 CKIAESNKD-VKIEPCGHLSCCLAAWQHSQTCFCRCBI 393
RESULT 11
ICP0_HSVEB
ID ICP0_HSVEB STANDARD; PRT; 532 AA.
AC P28950.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trans-acting transcriptional protein ICP0.
GN 63.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1";
RL Virology 189:304-316 (1992).
[2]
STRUCTURE BY NMR OF 1-63.
RX MEDLINE=94087718; PubMed=8263911;
RA Everett R.D., Barlow P.N., Milner A., Luisi B., Orr A., Hope G.,
RA Lyon D.;
RT "A novel arrangement of zinc-binding residues and secondary structure
in the C3HC4 motif of an alpha herpes virus protein family";
RL J. Mol. Biol. 234:1038-1047 (1993).

RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodargren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; BC021778; AAH21778.1; -
DR MGD; MGI:1929069; Rnf8.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Fing.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; RING; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Zinc-finger.
KW DOMAIN 38 92 FHA.
FT DOMAIN 279 345 GLN-RICH.
FT DOMAIN 406 444 RING-TYPE.
FT ZN FING 406 444
SQ SEQUENCE 488 AA; 55516 MW; 428242204EBC44A1 CRC64;
Query Match 40.8%; Score 108.5; DB 1; Length 488;
Best Local Similarity 45.0%; Pred. No. 2e-05;
Matches 18; Conservative 7; Mismatches 14; Indels 1; Gaps 1;
OY 2 CIIQDHLHDCVSLQPCVHHTFCAACYSQGMWERSLCTPCR 41
DB 406 CIIQSEFIEAVTLN-CAHSFCSFCINEMKRVCEPICR 444
RESULT 7
YUES CAEL STANDARD; PRT; 564 AA.
AC P90859; P90852;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical RING finger protein F26E4.11 in chromosome I.
GN F26E4.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RX [1]

SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC Lightning J., Baynes C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z81070; CAB03009.1; -
DR EMBL; Z81075; CAB03009.1; JOINED.
DR EMBL; Z81075; CAB03049.1; -
DR EMBL; Z81070; CAB03049.1; JOINED.
DR PIR; T21423; T21423.
DR WormPep; F26E4.11; CE09695.
DR InterPro; IPR003892; CUE.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF02845; CUE; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00546; CUE; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Zinc-finger.
FT ZN FING 335 373 RING-TYPE.
SQ SEQUENCE 564 AA; 64709 MW; 7C790C238207E49B CRC64;
Query Match 38.9%; Score 103.5; DB 1; Length 564;
Best Local Similarity 41.9%; Pred. No. 8.5e-05;
Matches 18; Conservative 6; Mismatches 18; Indels 1; Gaps 1;
OY 2 CIIQDHLHDCVSLQPCVHHTFCAACYSQGMWERSLCTPCR 44
DB 335 CVCWELLGTSERL-PCSHQHDWCLMWLAQDSSCTCRCTI 376
RESULT 8
BBAP HUMAN STANDARD; PRT; 740 AA.
AC Q8TD06;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE B-lymphoma- and BAL-associated protein (Rhysin 2) (Rhysin2).
GN BBAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., IN VITRO UBIQUITIN LIGASE ACTIVITY, AND
RP INTERACTION WITH BAL AND DTX1.
RX MEDLINE=22679154; PubMed=12670957;
RA Takeyama K., Aguilar R.C.T., Gu L., He C., Freeman G.J., Kutok J.L.,
RA Aster J.C., Shipp M.A.;
RT "the BAL-binding protein BBAP and related Deltex family members
RT exhibit ubiquitin-protein isopeptide ligase activity";
RL J. Biol. Chem. 278:21930-21937 (2003).
RN [2]
RN SEQUENCE FROM N.A.
RA Roberts R.C., Kendrick-Jones J., Jensen O.N.;
RT "rhysin2 is a novel protein identified by mass spectrometry found in a
RT myosin VI-containing complex isolated by immunoprecipitation";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Lymph;
RC MEDLINE=22388257; PubMed=12477932;
RX

```

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci F., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Villard M., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC
CC -|- SIMILARITY: Contains 1 FHA domain.
CC -|- SIMILARITY: Contains 1 RING-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB012770; BAA33557.1; -
DR EMBL; AB014546; BAA31621.1; -
DR EMBL; AL096712; CAB75689.1; -
DR EMBL; BC007517; AAH07517.1; -
DR Genew; HGNC:10071; RNFP8.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001841; Znf_Fing.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Zinc-finger.
FT DOMAIN 38 92 FHA.
FT DOMAIN 276 345 GLN-RICH.
FT ZN_FING 403 441 RING-TYPE.
FT SEQUENCE 485 AA; 55517 MW; 54650B2FFC9948B1 CRC64;
SQ
Query Match 40.8%; Score 108.5; DB 1; Length 485;
Best Local Similarity 45.0%; Pred. No. 2e-05;
Matches 18; Conservative 7; Mismatches 14; Indels 1; Gaps 1;
OY 2 CIIQDLHDCVSLQPCMTTFCACYSGWMSRSSLCPTCR 41
DB 403 CIIICSEYFIEAVTLN-CAHSFCSYCINEMWKRKIECPICR 441
RESULT 6
ID RNFS MOUSE
AC RNFS MOUSE STANDARD; PRT; 488 AA.
AC O8VC56;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RING finger protein 8.
DE RNFS.
GN RNFS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN RN
SEQUENCE FROM N.A.
PC TISSUE=Kidney;

```

100

Q06587 homo sapien
Q02398 emericella
Q9c035 homo sapien
P33288 neurospora
P23798 mus musculus
P35227 homo sapien
Q05568 saccharomyc
Q7y332 pan troglod
Q9udy6 homo sapien
Q9c029 homo sapien
P35820 drosophila
P38239 saccharomyc

34 89.5 33.6 377 1 RNGL HUMAN
35 88.5 33.3 443 1 UVSH EMENI
36 88.5 33.3 493 1 TRM5 HUMAN
37 88.5 33.3 501 1 UVS2 NEUCR
38 88 33.1 342 1 ME18 MOUSE
39 88 33.1 344 1 ME18 HUMAN
40 87.5 32.9 337 1 PEXA YEAST
41 87 32.7 481 1 RNF9 PANTR
42 87 32.7 482 1 RNF9 HUMAN
43 87 32.7 511 1 TRM7 HUMAN
44 87 32.7 1603 1 PSC DROME
45 85.5 32.1 115 1 YBR2 YEAST

ALIGNMENTS

RESULT 1
ICP0_VZVD
ID ICP0_VZVD STANDARD; PRT; 467 AA.
AC P09309;
DT 01-VAR-1989 (Rel. 10, Created)
DT 01-VAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-acting transcriptional protein ICP0.
GN 61.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
RL [2]
RP CHARACTERIZATION.
RX MEDLINE=93059681; PubMed=1366099;
RA Moriuchi H., Moriuchi M., Smith H.A., Straus S.E., Cohen J.I.;
RT "Varicella-zoster virus open reading frame 61 protein is functionally
homologous to herpes simplex virus type 1 ICP0.";
RL J. Virol. 66:7303-7308(1992).
CC -1- FUNCTION: REPRESSES THE EXPRESSION OF VIRAL IE, EARLY, AND LATE
GENE PROMOTERS.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ICP0 PROTEIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04370; CAA27944.1; --
CC FR; I27215; WZEE1.
CC HSP; P28990; ICHC.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00097; zf-C3HC4; 1.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC Transcription regulation; Trans-acting factor; Repressor; Zinc-finger;
CC DNA-binding.
CC ZN FING 19 58 RING-TYPE.
CC FT ZN FING 19 58 RING-TYPE.
CC SQ SEQUENCE 467 AA; 50916 MW; 25EFA6977EA6994C CRC64;

Query Match 48.5%; Score 129; DB 1; Length 467;
Best Local Similarity 45.5%; Pred. No. 8.9e-08;
Matches 20; Conservative 2; Mismatches 22; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:37:32 ; Search time 1.4868 Seconds
(without alignments)
1540.951 Million cell updates/sec

Title: US-10-048-046-2_COPY_303_346

Perfect score: 266
Sequence: 1 TCTICDILLHDCVSLQPCMH.....ACVSGHWERSLLCPTCRCPV 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	48.5	467	1 ICP0_VZVD	P09309 varicella-z
2	119	44.7	676	1 ICP0_HSVBJ	P29128 bovine herp
3	119	44.7	676	1 ICP0_HSVBK	P29128 bovine herp
4	110	41.4	610	1 YODA CAEL	Q09268 caenorhabdi
5	108.5	40.8	485	1 RN11 HUMAN	Q76064 homo sapien
6	108.5	40.8	488	1 RN11 HUMAN	Q76064 homo sapien
7	103.5	38.9	564	1 YFBS CAEL	Q8vc56 mus musculus
8	99	37.2	740	1 BBAP HUMAN	P08859 caenorhabdi
9	99	37.2	744	1 TRM3 HUMAN	O8tdb6 homo sapien
10	98	36.8	474	1 CBLC HUMAN	O75382 homo sapien
11	98	36.8	532	1 ICP0_HSVBJ	Q9ulv8 homo sapien
12	97	36.5	154	1 RN11 HUMAN	P28990 equine herp
13	97	36.5	154	1 RN11 HUMAN	Q9y3c5 homo sapien
14	97	36.5	825	1 ICP0_HSV2H	Q9y3c5 homo sapien
15	96	36.1	744	1 TRM3 MOUSE	Q9y3c5 mus musculus
16	96	36.1	744	1 TRM3 RAT	P28284 herpes simp
17	96	36.1	906	1 CBL HUMAN	O9rlr2 mus musculus
18	96	36.1	913	1 CBL MOUSE	O70277 rattus norv
19	96	36.1	982	1 CBL HUMAN	P22681 homo sapien
20	95.5	35.9	643	1 AMF2 HUMAN	P22682 mus musculus
21	95.5	35.9	643	1 AMF2 MOUSE	Q13191 homo sapien
22	95	35.7	410	1 ICP0_PRVIF	Q9ukv5 homo sapien
23	95	35.7	775	1 ICP0_HSV11	O9t049 mus musculus
24	94	35.3	324	1 BM11 FELCA	P29129 pseudorabie
25	94	35.3	326	1 BM11 FELCA	P25916 mus musculus
26	94	35.3	326	1 BM11 HUMAN	P25916 mus musculus
27	93.5	35.2	407	1 RP2 HUMAN	Q9est0 felis silve
28	93.5	35.2	407	1 RP2 HUMAN	P35226 homo sapien
29	92	34.8	488	1 RN23 MOUSE	O60858 homo sapien
30	92	34.6	518	1 RN23 HUMAN	O74747 schizosacch
31	91.5	34.4	605	1 RN23 HUMAN	Q9ean2 mus musculus
32	91.5	34.4	643	1 PJAI HUMAN	Q9hcm9 homo sapien
33	90.5	34.0	161	1 YODA CAEL	Q9hcm9 homo sapien

A; Introns: 88/3; 132/2; 174/3; 222/3; 334/1; 493/3

Query Match 38.9%; Score 103.5; DB 2; Length 564;
Best Local Similarity 41.9%; Pred. No. 0.0022;
Matches 18; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

QY 2 CIIQDQLHDCVSLQPCVHTFCAACYSYSGWMSRLCPTCRCPV 44
DB 335 CVCVCLNFKDDTLRLVPPPCVHVFDVCLWLAQDSSCPTCRCTI 376

RESULT 12
F71614
Chromatinic RING finger DRING protein homolog PF80440C - malaria parasite (Plasmodium fa
C; Species: Plasmodium falciparum
C; Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 02-Sep-2000
C; Accession: F71614
R; Gardner, M.J.; Tetelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Perteira, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A; Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A; Reference number: A71600; MUID: 99021743; PMID: 9804551
A; Accession: F71614
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Map position: 2
A; Molecule type: DNA
A; Residues: 1-568 <GAR>
A; Cross-references: GB:AE001395; GB:AE001362; NID:G3845184; PIDN:AAC71877.1; PID:G384518
A; Experimental source: clone 3D7
C; Genetics:
A; Gene: PF80440C
C; Superfamily: RING finger homology
F; 210-260/Domain: RING finger homology <RRN>

Query Match 38.0%; Score 101; DB 2; Length 568;
Best Local Similarity 37.8%; Pred. No. 0.004;
Matches 17; Conservative 11; Mismatches 15; Indels 2; Gaps 1;

QY 2 CIIQDQLHDCVSLQPCVHTFCAACYSYSGWMSRL--CPTCRCPV 44
DB 214 CPTCLSLIHEPVTLSFCHFWKCLATAIQKSIDNCPSCKTKI 258

RESULT 13
B85092
Hypothetical protein ATG09100 [imported] - Arabidopsis thaliana
C; Species: Arabidopsis thaliana (mouse-ear cress)
C; Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C; Accession: B85092
R; anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A; Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A; Reference number: A85001; MUID: 20083488; PMID: 10617198
A; Accession: B85092
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-132 <STO>
A; Cross-references: GB:NC_001268; NID:G7267553; PIDN:CAB78034.1; GSPDB:GN00140
C; Genetics:
A; Gene: ATG09100
A; Map position: 4

Query Match 37.8%; Score 100.5; DB 2; Length 132;
Best Local Similarity 39.1%; Pred. No. 0.0016;
Matches 18; Conservative 7; Mismatches 18; Indels 3; Gaps 1;

QY 2 CIIQDQLHDCVSLQ---PCVHTFCAACYSYSGWMSRLCPTCRCPV 44
DB 85 CVCVCLNFKDDTLRLVPPPCVHVFDVCLWLAQDSSCPTCRCTI 130

RESULT 14
T39456
zinc finger protein - fission yeast (Schizosaccharomyces pombe)

C; Species: Schizosaccharomyces pombe
C; Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 23-Sep-2002
C; Accession: T39456
R; Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Churcher, C.M.
submitted to the EMBL Data Library, June 1998
A; Reference number: 221855
A; Accession: T39456
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-486 <LYN>
A; Cross-references: EMBL:AL023780; NID:el295812; PIDN:CAA19328.1; GSPDB:GN00067; SPDB:SP
A; Experimental source: strain 972h-; cosmid c14F5
C; Genetics:
A; Gene: SPDB:SPBCL4F5.10C
A; Map position: 2
C; Superfamily: zinc finger protein; RING finger homology
P; 165-212/Domain: RING finger homology <RRN>

Query Match 37.8%; Score 100.5; DB 2; Length 486;
Best Local Similarity 47.5%; Pred. No. 0.004;
Matches 19; Conservative 5; Mismatches 15; Indels 1; Gaps 1;

QY 2 CIIQDQLHDCVSLQPCVHTFCAACYSYSGWMSRLCPTCR 41
DB 169 CQICFGMLYDPV-VSPCGHTFGPCLMQALTQSPQCPTCR 207

RESULT 15
B96674
Hypothetical protein F16G16.4 [imported] - Arabidopsis thaliana
C; Species: Arabidopsis thaliana (mouse-ear cress)
C; Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C; Accession: B96674
R; Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A; Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A; Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A; Reference number: A86141; MUID: 21016719; PMID: 11130712
A; Accession: B96674
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-496 <STO>
A; Cross-references: GB:AE005173; NID:G6227002; PIDN:AAF06038.1; GSPDB:GN00141
C; Genetics:
A; Gene: F16G16.4
A; Map position: 1

Query Match 37.8%; Score 100.5; DB 2; Length 496;
Best Local Similarity 43.2%; Pred. No. 0.0041;
Matches 19; Conservative 5; Mismatches 19; Indels 1; Gaps 1;

QY 1 TCIIQDQLHDCVSLQPCVHTFCAACYSYSGWMSRLCPTCRCPV 44
DB 327 TCIICREWTSAKKL-VGHLFHVHCLRSWLERQNTCTCRALV 369

Search completed: May 7, 2004, 14:50:08
Job time: 3.71251 secs

A:Molecule type: DNA
 A:Residues: 1-551 <HUG>
 A:Cross-references: EMBL:Z74755; NID:G1419784; PID:e252259; PID:G1419785; MIPS:YOL013c
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:HRD1
 A:Cross-references: SGD:S0005373; MIPS:YOL013c
 A:Map position: 15L
 A:Superfamily: RING finger homology
 C:Keywords: transmembrane protein
 F:10-26/Domain: transmembrane #status predicted <TM1>
 F:46-62/Domain: transmembrane #status predicted <TM2>
 F:105-121/Domain: transmembrane #status predicted <TM3>
 F:144-160/Domain: transmembrane #status predicted <TM4>
 F:185-201/Domain: transmembrane #status predicted <TM5>
 F:345-405/Domain: RING finger homology <RRN>
 Query Match 40.2%; Score 107; DB 2; Length 551;
 Best Local Similarity 40.0%; Pred. No. 0.00097;
 Matches 22; Conservative 2; Mismatches 19; Indels 12; Gaps 1;
 QY 2 CIIICQDLHDCVSLQ-----PCMHFFCAACYSGWMERSLLCPTCRCPV 44
 Db 349 CIIICWDELHSPNQTKWKNKKPKRLPCGHILSLCKNWMERSQTCPIRLPV 403
 RESULT 8
 F84591
 hypothetical protein At2g20650 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84591
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.A.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: F84591
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-276 <STO>
 A:Cross-references: GB:AB002093; NID:G4512649; PIDN:AAD21704.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g20650
 A:Map position: 2
 Query Match 39.7%; Score 105.5; DB 2; Length 276;
 Best Local Similarity 42.9%; Pred. No. 0.00086;
 Matches 21; Conservative 6; Mismatches 15; Indels 7; Gaps 3;
 QY 2 CIIICQ---DLAH---DCVSLQPCMHFFCAACYSGWMERSLLCPTCRCPV 44
 Db 226 CVICMTTIDLRINDCM-VTPCEHIFHSGCLQRMWDIKVECTCRRL 273
 RESULT 9
 T22687
 hypothetical protein F55A11.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
 C:Accession: T22687
 R:Kershaw, J.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19600
 A:Accession: T22687
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-610 <WIL>
 A:Cross-references: EMBL:Z72511; PIDN:CAA96657.1; GSPDB:GN00023; CESP:F55A11.3
 A:Experimental source: clone F55A11
 C:Genetics:
 A:Gene: CESP:F55A11.3

A:Map position: 5
 A:Introns: 76/3; 221/1; 275/2; 338/1; 424/3
 A:Superfamily: RING finger homology
 F:288-338/Domain: RING finger homology <RRN>
 Query Match 39.1%; Score 104; DB 2; Length 610;
 Best Local Similarity 41.3%; Pred. No. 0.0021;
 Matches 19; Conservative 6; Mismatches 19; Indels 2; Gaps 1;
 QY 1 TCIIICQDLHDCVSLQ---PCMHFFCAACYSGWMERSLLCPTCRCPV 44
 Db 291 TCIIICREMTVDASPKRLPCSHVFAHCLRSWFORQOTCPTCRTDI 336
 RESULT 10
 T16648
 hypothetical protein R02E12.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
 C:Accession: T16648
 R:Leimbach, D.
 submitted to the EMBL Data Library, April 1996
 A:Description: The sequence of C. elegans cosmid R02E12.
 A:Reference number: Z18554
 A:Accession: T16648
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-639 <LEI>
 A:Cross-references: EMBL:U53337; NID:G1255833; PIDN:AAA96184.1; GSPDB:GN000
 A:Experimental source: strain Bristol N2; clone R02E12
 C:Genetics:
 A:Gene: CESP:R02E12.4
 A:Map position: X
 A:Introns: 63/3; 90/1; 136/2; 177/1; 220/2; 242/3; 294/2; 427/2; 464/3; 505/3; 580/1
 C:Superfamily: RING finger homology
 F:22-70/Domain: RING finger homology <RRN>
 Query Match 39.1%; Score 104; DB 2; Length 639;
 Best Local Similarity 39.5%; Pred. No. 0.0022;
 Matches 17; Conservative 6; Mismatches 20; Indels 0; Gaps 0;
 QY 2 CIIICQDLHDCVSLQPCMHFFCAACYSGWMERSLLCPTCRCPV 44
 Db 26 CHICQVNHPEVTFITCKRHSICAGCAGRWLSSCSVCPMCRTV 68
 RESULT 11
 T21423
 hypothetical protein F26E4.11 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T21423; T21749
 R:Lightning, J.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19419
 A:Accession: T21423
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-564 <WIL>
 A:Cross-references: EMBL:Z81070; PIDN:CAB03009.1; GSPDB:GN00019; CESP:F26E4.11
 A:Experimental source: clone F26E4
 R:Baynes, C.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19468
 A:Accession: T21749
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-564 <WIL>
 A:Cross-references: EMBL:Z81075; PIDN:CAB03049.1; GSPDB:GN00019; CESP:F26E4.11
 A:Experimental source: clone F35C12
 C:Genetics:
 A:Gene: CESP:F26E4.11
 A:Map position: 1

```

RESULT 5
T15741
hypothetical protein C32D5.10 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C/Accession: T15741
R/Miller, N.
submitted to the EMBL Data Library, July 1995
A/Description: The sequence of C. elegans cosmid C32D5.
A/Reference number: Z18396
A/Accession: T15741
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-621 <MIL>
A/Cross-references: EMBL:U25511; NID:G746463; PID:G746473; PIDN:AAC46798.1;
A/Experimental_source: strain Bristol N2
C/Genetics:
A/Gene: CBSP:C32D5.10
A/Introns: 47/2; 141/2; 248/3; 348/1; 383/1; 406/3; 429/1; 592/2; 612/1
C/Suprafamily: RING finger homology
F;37-87/Domain: RING finger homology <RRN>

Query Match 41.4%; Score 110; DB 2; Length 621;
Best Local Similarity 40.0%; Pred. No. 0.00053;
Matches 18; Conservative 7; Mismatches 18; Indels 2; Gaps

QY 2 CIIICDPLHDCVSLQPCMHFTCAACYSGWMMRSS--LCPTCRCPV 44
      : : : : : : : : : : : : : : : : : : : : : : : :
Db 41 CSVCKNEIIDTWSLSDCCHEFCYDCIVGWLTKSGSPFCPCMKTFV 85

RESULT 6
T39702
probable peroxisome assembly protein - fission yeast (Schizosaccharomyces F
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C/Accession: T39702
R/Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, July 1999
A/Reference number: Z21870
A/Accession: T39702
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1306 <WOO>
A/Cross-references: EMBL:AL109652; PIDN:CAE51769.1; GSPDB:GN00067
A/Experimental_source: strain 972h-; cosmid c17A3
C/Genetics:
A/Gene: pi037
A/Map position: 2
A/Introns: 13/1
C/Suprafamily: RING finger homology
F;252-259/Domain: RING finger homology <RRN>

Query Match 40.4%; Score 107.5; DB 2; Length 306;
Best Local Similarity 40.0%; Pred. No. 0.00058;
Matches 16; Conservative 7; Mismatches 16; Indels 1; Gaps

QY 2 CIIICDPLHDCVSLQPCMHFTCAACYSGWMMRSSLCPTCR 41
      : : : : : : : : : : : : : : : : : : : : : :
Db 256 CSLCMFPIH-CPAATCGHPCWSCINGWTSKKSCTCLCR 294

RESULT 7
S66595
probable membrane protein YOL013c - Yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein O2341
C/Species: Saccharomyces cerevisiae
C/Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 02-Sep-2000
C/Accession: S66595
R/Hughes, B.; Pohl, T.M.
submitted to the Protein Sequence Database, July 1996
A/Reference number: S66585
A/Accession: S66595

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:41:17 ; Search time 2.46251 Seconds
(without alignments)
1718.743 Million cell updates/sec

Title: US-10-048-046-2_COPY_303_346

Perfect score: 266
Sequence: 1 TCIIICDLDHDCVSLQPCVH.....ACVSGWMERSLLCPTCRCPV 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*

- 1: pir1:*
- 2: pir2:*
- 3: pir3:*
- 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135	50.8	473	2 C96516	Fl6N3.15 [imported]
2	129	48.5	467	1 WZBE61	gene 61 protein -
3	119	44.7	676	1 EDBE22	immediate-early pr
4	119	44.7	676	1 EDBE23	immediate-early pr
5	110	41.4	621	2 T15741	hypothetical prote
6	107.5	40.4	306	2 T39702	probable peroxisom
7	107	40.2	551	2 S66695	probable membrane
8	105.5	39.7	276	2 P84591	hypothetical prote
9	104	39.1	610	2 T22687	hypothetical prote
10	104	39.1	639	2 T16648	hypothetical prote
11	103.5	38.9	564	2 T21423	hypothetical prote
12	101	38.0	568	2 P71614	chromatinic RING f
13	100.5	37.8	132	2 B5092	hypothetical prote
14	100.5	37.8	486	2 T39456	zinc finger protei
15	100.5	37.8	496	2 B96674	hypothetical prote
16	100.5	37.8	1208	2 T00362	hypothetical prote
17	100	37.6	222	2 JC4296	ring finger protei
18	99.5	37.4	1375	2 T37672	probable DNA repai
19	98	36.8	230	2 JC7972	spermatogenesis-re
20	98	36.8	532	1 WZBEF5	59K transcription
21	98	36.8	536	2 T42606	probable transcrip
22	97	36.5	410	2 A40505	early protein EF0
23	97	36.5	825	1 EDBEXD	immediate-early pr
24	96.5	36.3	409	2 T25935	hypothetical prote
25	96	36.1	906	2 A43817	transforming prote
26	95.5	35.9	245	2 T45652	RNA binding-like p
27	95	35.7	315	2 D48560	immediate-early pr
28	95	35.7	775	1 EDBE11	immediate-early pr
29	94	35.3	118	2 T01883	hypothetical prote

30	94	35.3	324	2 A39523	transforming prote
31	94	35.3	326	2 I51694	transforming prote
32	94	35.3	326	2 I54339	proto-oncogene bml
33	94	35.3	532	2 T49467	related to COP1-in
34	93.5	35.2	242	2 T48173	hypothetical prote
35	93.5	35.2	1331	2 T04938	hypothetical prote
36	93.5	35.2	1495	2 A85240	hypothetical prote
37	93.5	35.2	1495	2 T10649	hypothetical prote
38	92.5	34.8	387	2 T39653	probable DNA repai
39	92	34.6	273	2 T47692	testis-abundant fi
40	92	34.6	518	2 JC7387	hypothetical prote
41	91.5	34.4	302	2 C85092	hypothetical prote
42	91.5	34.4	327	2 D86474	probable RING zinc
43	91.5	34.4	345	2 D85092	hypothetical prote
44	91.5	34.4	836	2 T18460	hypothetical prote
45	91.5	34.4	2447	2 T16870	hypothetical prote

ALIGNMENTS

RESULT 1
C96516
Fl6N3.15 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96516
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. A:Authors: Salzberg, S.L.; Schwartz, J.R.; Venter, J.C.; Davis, R.W. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C96516
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <STO>
A:Cross-references: GB:AE005173; NID:G5668816; PIDN:AA046042.1; GSPDB:GN00141
C:Genetics:
A:Gene: Fl6N3.15
A:Map position: 1

Query Match 50.8%; Score 135; DB 2; Length 473;
Best Local Similarity 46.9%; Pred. No. 1.3e-06;
Matches 23; Conservative 5; Mismatches 15; Indels 6; Gaps 1;

QY 2 CIIICDLDHDCVSLQPCVHFCACVSGWMERS-----LCPTCRCPV 44
Db 145 CSICLNWHDVVTAPCLHFNFCGSEWMRSEKHKHVLCPQCRTTV 193

RESULT 2
WZBE61
gene 61 protein - human herpesvirus 3
C:Species: human herpesvirus 3, varicella-zoster virus
C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 17-Mar-2000
C:Accession: I27215
J. Gen. Virol. 67, 1759-1816, 1986
R:Davidson, A.J.; Scott, J.E.
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657; PMID:3018124
A:Accession: I27215
A:Molecule type: DNA
A:Residues: 1-467 <DAV>
A:Cross-references: EMBL:X04370; NID:G59989; PIDN:CAA27944.1; PID:G60050
C:Genetics:
A:Gene: 61
C:Superfamily: varicella-zoster virus gene 61 protein; RING finger homology

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (358)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (494)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (504)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (505)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-808

```

```

Query Match      38.7%; Score 103; DB 11; Length 505;
Best Local Similarity 39.5%; Pred. No. 0.0035;
Matches 17; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

```

```

QY      2 CIIICODLLHDCVSLQPCMHFTCAACYSGWMSRSSLCPTCRCPV 44
DB      291 CIIICREEMVTGAKRLPCNHFHTSCLRSWFORQQTCTCRMDV 333

```

```

RESULT 15
US-09-833-245-810
; Sequence 810, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 810
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-810

```

```

Query Match      38.7%; Score 103; DB 11; Length 617;
Best Local Similarity 39.5%; Pred. No. 0.0041;
Matches 17; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

```

```

QY      2 CIIICODLLHDCVSLQPCMHFTCAACYSGWMSRSSLCPTCRCPV 44
DB      291 CIIICREEMVTGAKRLPCNHFHTSCLRSWFORQQTCTCRMDV 333

```

```

Search completed: May 7, 2004, 15:06:51
Job time : 7.96938 secs

```


US-10-424-599-277916
 ; Sequence 277916, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 277916
 ; LENGTH: 257
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_92981C.1.pap
 ; US-10-424-599-277916

Query Match 42.3%; Score 112.5; DB 12; Length 257;
 Best Local Similarity 46.5%; Pred. No. 0.00018;
 Matches 20; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 CIIICODLLHDCVSLQPCMHHTFCAACYSYSGWMSRLCPTCRCPV 44
 DB 196 CAICQKMHAPILLR-CXHFCDVCSEWFERTCPLCALV 237

RESULT 9
 US-10-425-114-51122
 ; Sequence 51122, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 51122
 ; LENGTH: 327
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700083617_FLI.pap
 ; US-10-425-114-51122

Query Match 42.3%; Score 112.5; DB 12; Length 327;
 Best Local Similarity 48.8%; Pred. No. 0.00023;
 Matches 21; Conservative 3; Mismatches 18; Indels 1; Gaps 1;

QY 2 CIIICODLLHDCVSLQPCMHHTFCAACYSYSGWMSRLCPTCRCPV 44
 DB 266 CAICQKMHAPILLR-CXHFCDVCSEWFERTCPLCALV 307

RESULT 10
 US-10-424-599-235446
 ; Sequence 235446, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua

RESULT 6
 US-10-001-885-150
 ; Sequence 150, Application US/10001885
 ; Publication No. US20040058319A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Salceda, Susana
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Recipon, Herve
 ; APPLICANT: Cafferkey, Robert
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Chenghua
 ; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
 ; FILE REFERENCE: DEX-0279
 ; CURRENT APPLICATION NUMBER: US/10/001,885
 ; CURRENT FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 60/252,061
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: 60/253,257
 ; PRIOR FILING DATE: 2000-11-27
 ; NUMBER OF SEQ ID NOS: 167
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 150
 ; LENGTH: 254
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-10-001-885-150

Query Match 43.2%; Score 115; DB 12; Length 254;
 Best Local Similarity 41.5%; Pred. No. 9.8e-05;
 Matches 17; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 TCIIICODLLHDCVSLQPCMHHTFCAACYSYSGWMSRLCPTCR 41
 DB 21 TCRLCSGLIDATTTCLEHTFCRSLVLYLENNCTCTCR 61

RESULT 7
 US-10-425-114-40033
 ; Sequence 40033, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 40033
 ; LENGTH: 183
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700962291_FLI.pap
 ; US-10-425-114-40033

Query Match 42.3%; Score 112.5; DB 12; Length 183;
 Best Local Similarity 46.5%; Pred. No. 0.00014;
 Matches 20; Conservative 4; Mismatches 18; Indels 1; Gaps 1;

QY 2 CIIICODLLHDCVSLQPCMHHTFCAACYSYSGWMSRLCPTCRCPV 44
 DB 122 CAICQKMHAPILLR-CXHFCDVCSEWFERTCPLCALV 163

RESULT 8

```

; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 809
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (414)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (415)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (420)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-809

Query Match
Best Local Similarity 100.0%; Score 266; DB 9; Length 426;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCICQDLHDCVSLQPCMTFCAACYSYGMWERSLCTCRCPV 44
Db 213 TCICQDLHDCVSLQPCMTFCAACYSYGMWERSLCTCRCPV 256

RESULT 3
US-09-780-525-2
; Sequence 2, Application US/09780525
; Patent No. US2002004223A1
; GENERAL INFORMATION:
; APPLICANT: Bin-Bing Zhou
; APPLICANT: Yuan Zhu
; APPLICANT: Priya Chaturvedi
; APPLICANT: Mark R. Hurie
; APPLICANT: Xiaotong Li
; TITLE OF INVENTION: FHAR1, A NEW RING FINGER PROTEIN
; FILE REFERENCE: GP-70668-C1
; CURRENT APPLICATION NUMBER: US/09/780,525
; CURRENT FILING DATE: 2001-02-09
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 664
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-780-525-2

Query Match
Best Local Similarity 100.0%; Score 266; DB 9; Length 664;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCICQDLHDCVSLQPCMTFCAACYSYGMWERSLCTCRCPV 44
Db 303 TCICQDLHDCVSLQPCMTFCAACYSYGMWERSLCTCRCPV 346

RESULT 4
US-09-764-864-1270
; Sequence 1270, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 1270
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (189)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (192)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (193)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (194)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (213)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (223)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (229)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (230)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1270

Query Match
Best Local Similarity 100.0%; Score 176; DB 9; Length 230;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 POMTFCAACYSYGMWERSLCTCRCPV 44
Db 5 POMTFCAACYSYGMWERSLCTCRCPV 32

RESULT 5
US-09-919-039-285
; Sequence 285, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 285
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1427459CD1
US-09-919-039-285

Query Match
Best Local Similarity 43.2%; Score 115; DB 10; Length 247;
Matches 17; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 TCICQDLHDCVSLQPCMTFCAACYSYGMWERSLCTCR 41
Db 21 TCICQDLHDCVSLQPCMTFCAACYSYGMWERSLCTCR 61

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:49:13 ; Search time 6.96938 Seconds
(without alignments)
1752.371 Million cell updates/sec

Title: US-10-048-046-2_COPY_303_346
Perfect score: 266
Sequence: 1 TCIIQDLHDCVSLQPCMH.....ACYSGMWSSSLCPTCRCPV 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	266	100.0	349	15	US-10-108-260A-4516
2	266	100.0	426	9	US-09-764-864-809
3	266	100.0	684	9	US-09-780-525-2
4	176	66.2	230	9	US-09-764-864-1270
5	115	43.2	247	10	US-09-919-039-285
6	115	43.2	254	12	US-10-001-885-150
7	112.5	42.3	183	12	US-10-425-114-40033
8	112.5	42.3	257	12	US-10-424-599-277916
9	112.5	42.3	327	12	US-10-425-114-51122
10	109.5	41.2	216	12	US-10-424-599-235446
11	108.5	40.8	242	12	US-10-424-599-250642
12	104.5	39.3	234	12	US-10-425-114-49876
13	104.5	39.3	393	12	US-10-424-599-154629
14	103	38.7	505	11	US-09-833-245-808
15	103	38.7	617	11	US-09-833-245-810

16	103	38.7	617	12	US-10-470-360-17	Sequence 17, Appl
17	100.5	37.8	185	12	US-10-425-114-40364	Sequence 40364, A
18	99.5	37.4	165	12	US-10-424-599-170452	Sequence 170452,
19	99.5	37.4	169	12	US-10-424-599-170451	Sequence 170451,
20	99.5	37.4	176	12	US-10-425-114-49027	Sequence 49027, A
21	99.5	37.4	206	12	US-10-424-599-180706	Sequence 180706,
22	99	37.2	163	9	US-09-764-864-845	Sequence 845, App
23	99	37.2	218	9	US-09-764-864-1300	Sequence 1300, App
24	99	37.2	218	14	US-10-106-698-6052	Sequence 6052, Ap
25	99	37.2	356	15	US-10-136-728-18	Sequence 18, Appl
26	99	37.2	356	15	US-10-094-749-1721	Sequence 1721, Ap
27	99	37.2	568	12	US-10-126-103-112	Sequence 112, App
28	99	37.2	739	9	US-09-957-635-2	Sequence 2, Appl
29	98.5	37.0	48	14	US-10-283-500-14	Sequence 14, Appl
30	98.5	37.0	113	12	US-10-424-599-158001	Sequence 158001,
31	98.5	37.0	486	12	US-10-425-114-40759	Sequence 40759, A
32	98.5	37.0	575	12	US-10-424-599-157327	Sequence 157327,
33	98	36.8	131	15	US-10-264-049-2856	Sequence 2856, Ap
34	98	36.8	158	9	US-09-764-864-1304	Sequence 1304, App
35	98	36.8	230	15	US-10-395-027-702	Sequence 702, App
36	98	36.8	439	9	US-09-764-864-849	Sequence 849, App
37	98	36.8	439	15	US-10-264-049-1081	Sequence 1081, Ap
38	98	36.8	474	14	US-10-157-031-251	Sequence 251, App
39	97.5	36.7	331	12	US-10-425-114-37376	Sequence 37376, A
40	97.5	36.7	353	12	US-10-424-599-187336	Sequence 187336,
41	97.5	36.7	420	12	US-10-424-599-160966	Sequence 160966,
42	97.5	36.7	424	12	US-10-425-114-70107	Sequence 70107, A
43	97	36.5	154	12	US-10-231-625-105	Sequence 105, App
44	97	36.5	207	9	US-09-925-302-638	Sequence 638, App
45	97	36.5	207	12	US-09-925-302-638	Sequence 638, App

ALIGNMENTS

RESULT 1
US-10-108-260A-4516
; Sequence 4516, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1 full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4516
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-108-260A-4516

Query Match 100.0%; Score 266; DB 15; Length 349;
Best Local Similarity 100.0%; Pred. No. 8.2e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCIIQDLHDCVSLQPCMHITCAACYSGMWSSSLCPTCRCPV 44
DB 103 TCIIQDLHDCVSLQPCMHITCAACYSGMWSSSLCPTCRCPV 146

RESULT 2
US-09-764-864-809
; Sequence 809, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper

FILING DATE: 31-Mar-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-198 CIP 1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-052-089A-12

Query Match 36.1%; Score 96; DB 4; Length 47;
Best Local Similarity 38.6%; Pred. No. 0.00032;
Matches 17; Conservative 6; Mismatches 19; Indels 2; Gaps 2;

QY 2 CIIICODLLHDCVSLQPCNHTFCACYSGWMMERSL-CPTCRCPV 44
DB 4 CKICAEKND-VKIEPCGHLMTCTSLTWSOESGCGCFRCBEI 46

RESULT 14
US-07-945-283-6
Sequence 6, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
TITLE OF INVENTION: Involving the EP0 and LIT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P.
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Herpes simplex virus
US-07-945-283-6

Query Match 35.7%; Score 95; DB 1; Length 69;
Best Local Similarity 39.0%; Pred. No. 0.00058;
Matches 16; Conservative 4; Mismatches 19; Indels 2; Gaps 1;

QY 2 CIIICODLLHDCVSLQPCNHTFCACYSGWMMERSL-CPTC 40
DB 6 CAVCTDEIAPHLRCDTFPCNHRFCIPCKMTWQIRNTCPIC 46

RESULT 15
US-07-945-283-4
Sequence 4, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
TITLE OF INVENTION: Involving The EP0 and LIT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P.
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-945-283-4

Query Match 35.7%; Score 95; DB 1; Length 410;
Best Local Similarity 39.5%; Pred. No. 0.003;
Matches 17; Conservative 2; Mismatches 24; Indels 0; Gaps 0;

QY 2 CIIICODLLHDCVSLQPCNHTFCACYSGWMMERSL-CPTCRCPV 44
DB 46 CPICLDVAATEAQTLPCNHRFCICIQRWTLTSTACPLCNARV 88

Search completed: May 7, 2004, 14:51:23
Job time : 3.92714 secs

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06639
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250858
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/331394
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 897P2PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-06639-14

Query Match 37.0%; Score 98.5; DB 5; Length 48;
Best Local Similarity 42.2%; Pred. No. 0.00017;
Matches 19; Conservative 5; Mismatches 16; Indels 5; Gaps 2;
QY 2 CIIICDILLHDCVSLQPCMTFCAACVSGWMMER-SSLCPTCRC 42
DB 4 CKICAENDKD-VKIEFCGLMCTSLTWSQSEGGSGGCPFCR 47

RESULT 11
US-07-945-283-5
Sequence 5, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudorabies virus Deletion Mutants
TITLE OF INVENTION: Involving The EPO and LIT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSES: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P.
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext.513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 67 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Pseudorabies virus
US-07-945-283-5
Query Match 36.5%; Score 97; DB 1; Length 67;
Best Local Similarity 39.5%; Pred. No. 0.00034;
Matches 17; Conservative 3; Mismatches 23; Indels 0; Gaps 0;
QY 2 CIIICDILLHDCVSLQPCMTFCAACVSGWMMERSSLCPTCRCPV 44
DB 6 CPICLDVAATEAQTLPCHKEFCLDIOIOWTLTSTACPLCKARV 48
RESULT 12
US-09-894-998A-47
Sequence 47, Application US/09894998A
Patent No. 6537555
GENERAL INFORMATION:
APPLICANT: Hosken, Nancy Ann
APPLICANT: Craig H. Day
APPLICANT: Davin C. Dillon
APPLICANT: McGowan, Patrick
APPLICANT: Sleath, Paul R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
FILE REFERENCE: 210121.538
CURRENT APPLICATION NUMBER: US/09/894,998A
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 47
LENGTH: 826
TYPE: PRT
ORGANISM: HSV-2
US-09-894-998A-47

Query Match 36.5%; Score 97; DB 4; Length 826;
Best Local Similarity 35.6%; Pred. No. 0.0034;
Matches 16; Conservative 6; Mismatches 21; Indels 2; Gaps 1;
QY 2 CIIICDILLHDCVSLQ-PCMTFCAACVSGWMMERSSLCPTCRCPV 44
DB 126 CAVCTDEIAPPLRCQSFCLHPFCIPCKMTWIPURNTCPLCNTFV 170
RESULT 13
US-09-052-089A-12
Sequence 12, Application US/09052089A
Patent No. 6346605
GENERAL INFORMATION:
APPLICANT: Lee, Soo Y.
Choi, Yongwon
TITLE OF INVENTION: SIGNAL TRANSDUCER FOR THE TNF RECEPTOR SUPER
FAMILY, AND USES THEREOF
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/052,089A

FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/331394
FILING DATE: 28-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 897P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-446-915-14

Query Match 37.0%; Score 98.5; DB 1; Length 48;
Best Local Similarity 42.2%; Pred. No. 0.00017;
Matches 19; Conservative 5; Mismatches 16; Indels 2;

Qy 2 CIIICODLLHDCVSLQPCMHFTCAACYSGWMER----SSLCPTRC 42
Db 4 CKICAENDKD-VKIEPCGHLMTCLTSWQSEGGSGGCGPCRC 47

RESULT 8
US-08-744-139-14
Sequence 14, Application US/08744139
Patent No. 5869612
GENERAL INFORMATION:
APPLICANT: Goeddel, David V
APPLICANT: Rothe, Mike
TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR-ASSOCIATED FACTORS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
FILING DATE: 31-Oct-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/250858
FILING DATE: 05/27/1994
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P0897C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-744-139-14

Query Match 37.0%; Score 98.5; DB 2; Length 48;
Best Local Similarity 42.2%; Pred. No. 0.00017;

Matches 19; Conservative 5; Mismatches 16; Indels 2;

Qy 2 CIIICODLLHDCVSLQPCMHFTCAACYSGWMER----SSLCPTRC 42
Db 4 CKICAENDKD-VKIEPCGHLMTCLTSWQSEGGSGGCGPCRC 47

RESULT 9
US-08-779-599-14
Sequence 14, Application US/08779599
Patent No. 6500922
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.
APPLICANT: Rothe, Mike
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
FILING DATE: 07-Jan-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P0897C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-779-599-14

Query Match 37.0%; Score 98.5; DB 4; Length 48;
Best Local Similarity 42.2%; Pred. No. 0.00017;
Matches 19; Conservative 5; Mismatches 16; Indels 2;

Qy 2 CIIICODLLHDCVSLQPCMHFTCAACYSGWMER----SSLCPTRC 42
Db 4 CKICAENDKD-VKIEPCGHLMTCLTSWQSEGGSGGCGPCRC 47

RESULT 10
PCT-US95-06639-14
Sequence 14, Application PC/TUS9506639
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb disk
COMPUTER: IBM PC compatible

APPLICANT: Rothe, Mike
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
CITY: 460 Point San Bruno Blvd
STATE: South San Francisco
COUNTRY: California
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/250.858
FILING DATE: 27-May-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dregler, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 897.1
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-250-858-14

Query Match 37.0%; Score 98.5; DB 1; Length 48;
Best Local Similarity 42.2%; Pred. No. 0.00017; Indels 5; Gaps 2;
Matches 19; Conservative 5; Mismatches 16;

QY 2 CIIQDLHDCVSLQPCMHFTFCAACYSQWMSRSLCPTCPC 42
DB 4 CKICAENDKD-VKIEPCGHLMTCLTSQWSEGGSGCPCPC 47

RESULT 7
US-08-446-915-14
Sequence 14, Application US/08446915
Patent No. 5741667
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.
APPLICANT: Rothe, Mike
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
CITY: 460 Point San Bruno Blvd
STATE: South San Francisco
COUNTRY: California
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,915
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250858

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
US-09-628-133-25

Query Match 40.2%; Score 107; DB 4; Length 551;
Best Local Similarity 40.0%; Pred. No. 0.00018; Indels 12; Gaps 1;
Matches 22; Conservative 2; Mismatches 19;

QY 2 CIIQDLHDCVSLQPCMHFTFCAACYSQWMSRSLCPTCPC 44
DB 349 CIIQDLHDCVSLQPCMHFTFCAACYSQWMSRSLCPTCPC 403

RESULT 5
US-08-331-394-14
Sequence 14, Application US/08331394
Patent No. 5670319
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.
APPLICANT: Rothe, Mike
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
CITY: 460 Point San Bruno Blvd
STATE: South San Francisco
COUNTRY: California
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331.394
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250858
FILING DATE: 27-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Dregler, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 897P1
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-331-394-14

Query Match 37.0%; Score 98.5; DB 1; Length 48;
Best Local Similarity 42.2%; Pred. No. 0.00017; Indels 5; Gaps 2;
Matches 19; Conservative 5; Mismatches 16;

QY 2 CIIQDLHDCVSLQPCMHFTFCAACYSQWMSRSLCPTCPC 42
DB 4 CKICAENDKD-VKIEPCGHLMTCLTSQWSEGGSGCPCPC 47

RESULT 6
US-08-250-858-14
Sequence 14, Application US/08250858
Patent No. 5708142
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.

RESULT 2

US-08-699-103B-25
; Sequence 25, Application US/08699103B
; Patent No. 6107462
; GENERAL INFORMATION:
; APPLICANT: Rine, Jasper D.
; APPLICANT: Hampton, Randolph
; TITLE OF INVENTION: GENES AND PROTEINS CONTROLLING
; TITLE OF INVENTION: CHOLESTEROL SYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/699.103B
; FILING DATE: 16-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002,581
; FILING DATE: 17-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 09272/005001
; TELEPHONE: 650/322-5070
; TELEFAX: 650/854-0875
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 551 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-08-699-103B-25

Query Match 40.2%; Score 107; DB 3; Length 551;
Best Local Similarity 40.0%; Pred.No.0.00018;
Matches 22; Conservative 2; Mismatches 19; Indels 12; Gaps 1;

QY 2 CIIICDILLHDCVSLQ-----PCMHFTCAACYSQWMSRSLCPTCRCPV 44
DB 349 CIIICWDELIHSPNQOTWKNKKPKRLPCGHILHLSCLKNNMERSQTCPCICRLPV 403

RESULT 3

US-09-229-059-25
; Sequence 25, Application US/09229059
; Patent No. 6333172
; GENERAL INFORMATION:
; APPLICANT: Rine, Jasper D.
; APPLICANT: Hampton, Randolph
; TITLE OF INVENTION: GENES AND PROTEINS CONTROLLING
; TITLE OF INVENTION: CHOLESTEROL SYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/229,059
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/699,103
FILING DATE: 16-AUG-1996
APPLICATION NUMBER: 60/002,581
FILING DATE: 17-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 09272/005001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/322-5070
TELEFAX: 650/854-0875
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
US-09-229-059-25

Query Match 40.2%; Score 107; DB 4; Length 551;
Best Local Similarity 40.0%; Pred.No.0.00018;
Matches 22; Conservative 2; Mismatches 19; Indels 12; Gaps 1;

QY 2 CIIICDILLHDCVSLQ-----PCMHFTCAACYSQWMSRSLCPTCRCPV 44
DB 349 CIIICWDELIHSPNQOTWKNKKPKRLPCGHILHLSCLKNNMERSQTCPCICRLPV 403

RESULT 4

US-09-628-133-25
; Sequence 25, Application US/09628133
; Patent No. 6531292
; GENERAL INFORMATION:
; APPLICANT: Rine, Jasper D.
; APPLICANT: Hampton, Randolph
; TITLE OF INVENTION: GENES AND PROTEINS CONTROLLING
; TITLE OF INVENTION: CHOLESTEROL SYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/628,133
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,103
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 09272/005001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/322-5070
; TELEFAX: 650/854-0875
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 551 amino acids

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: May 7, 2004, 14:42:23 ; Search time 2.92714 Seconds
(without alignments)
776.028 Million cell updates/sec

Title: US-10-048-046-2_COPY_303_346
 Perfect score: 266
 Sequence: 1 TCIIQCQLLHDCVSLQPCMH.....ACYSGWMERSLCPCTCRCPV 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues
Total number of hits satisfying chosen parameters: 389414

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database :
Issued Patents Ah.*
1: /cgn2_s/prodata/2/aaa/5a COMB pep.*
2: /cgn2_s/prodata/2/aaa/5b COMB pep.*
3: /cgn2_s/prodata/2/aaa/5c COMB pep.*
4: /cgn2_s/prodata/2/aaa/5d COMB pep.*
5: /cgn2_s/prodata/2/aaa/ACTUS COMB pep.*
6: /cgn2_s/prodata/2/aaa/backfile1 pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query *		Length	DB	ID	Description
		Match					
1	129	48.5	67	1	US-07-945-283-7	Sequence 7, Appli	
2	107	40.2	551	3	US-08-699-103B-25	Sequence 25, Appli	
3	107	40.2	551	4	US-09-329-059-25	Sequence 25, Appli	
4	107	40.2	551	4	US-09-628-133-25	Sequence 25, Appli	
5	98.5	37.0	48	1	US-08-331-394-14	Sequence 14, Appli	
6	98.5	37.0	48	1	US-08-250-858-14	Sequence 14, Appli	
7	98.5	37.0	48	1	US-08-446-915-14	Sequence 14, Appli	
8	98.5	37.0	48	2	US-08-744-139-14	Sequence 14, Appli	
9	98.5	37.0	48	4	US-08-779-599-14	Sequence 14, Appli	
10	98.5	37.0	48	5	PCT-US95-06639-14	Sequence 14, Appli	
11	97	36.5	67	1	US-07-945-283-5	Sequence 5, Appli	
12	97	36.5	826	4	US-09-894-998A-47	Sequence 47, Appli	
13	96	36.1	47	4	US-09-052-089A-12	Sequence 12, Appli	
14	95	35.7	69	1	US-07-945-283-6	Sequence 6, Appli	
15	95	35.7	410	1	US-07-945-283-4	Sequence 4, Appli	
16	94	35.3	326	3	US-08-786-606-8	Sequence 8, Appli	
17	92.5	34.8	336	3	US-08-972-719-2	Sequence 2, Appli	
18	92.5	34.8	336	3	US-08-753-038-2	Sequence 2, Appli	
19	92.5	34.8	326	4	US-09-150-175-2	Sequence 5, Appli	
20	91.5	34.4	223	4	US-09-327-983-5	Sequence 9, Appli	
21	91.5	34.4	395	2	US-08-841-349-9	Sequence 9, Appli	
22	91.5	34.4	395	4	US-09-431-184A-9	Sequence 36, Appli	
23	91	34.2	245	3	US-08-897-340-36	Sequence 36, Appli	
24	91	34.2	245	3	US-09-352-329-36	Sequence 8, Appli	
25	91	34.2	317	4	US-09-821-099A-8	Sequence 114, Appli	
26	90	33.8	41	1	US-09-599-360B-25	Sequence 84, Appli	
27	89.5	33.6	41	1	US-08-480-784-84		

ALIGNMENTS

RESULT 1

```

US-07-945-283-7
; Sequence 7, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EP0 and LLT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P.
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext.513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 67 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Varicella-zoster virus
US-07-945-283-7

```

```

Query Match      48.5%; Score 129; DB 1; Length 67;
Best Local Similarity 45.5%; Pred.No. 9.5e-08;
Matches 20; Conservative 2; Mismatch 22; Indels
QY 1 TCIIQQDILHDCVSLQPCMHHTCAACYSQWMBRSSLCPTCRCPV 44
DB 5 TCTICWSTVSDLGKTMPCILHDFCFVCIRAWSTSVQCPLCRCPV 48

```



```
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 50.8%; Score 135; DB 3; Length 453;
Best Local Similarity 46.9%; Pred. No. 1.2e-06;
Matches 23; Conservative 5; Mismatches 15; Indels 6; Gaps 1;

QY 2 CIIQDILLHDCVSLQPCNHTFCAACYSGWNERSS-----LCPTCRCPV 44
Db 125 CSICLNIWHDVVTAAAPCLHNFNGCFSEWWRSEKHKHYLCPCQRTTV 173
```

Search completed: May 7, 2004, 14:45:48
Job time : 11.2637 secs

PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 05-MAY-1999; 99US-0132407P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139464P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139751P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149175P.
PR 17-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149928P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.

[REDACTED]

```
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145132P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147303P.
PR 08-AUG-1999; 99US-0147493P.
PR 08-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156599P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 50.8%; Score 135; DB 3; Length 453;
Best Local Similarity 46.9%; Pred. No. 1.2e-06;
Matches 23; Conservative 5; Mismatches 15; Indels 6; Gaps 1;

QY 2 CIIICODLLHDCVSLQPCMHFTCAACYSQWMSRSS-----LCPTCRCPV 44
DB 125 CSICLINIWDVVTAAPCLHNFQNGCFSEWMSRSEKHKHVLCPQCRITV 173

RESULT 15
AAG37658
ID AAG37658 standard; protein; 453 AA.
XX
AC AAG37658;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46342.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
```

PR	31-AUG-1999;	99US-0151438P.	PN	EP1033405-A2.
PR	01-SEP-1999;	99US-0151930P.	XX	
PR	07-SEP-1999;	99US-0152363P.	PD	06-SEP-2000.
PR	10-SEP-1999;	99US-0153070P.	XX	
PR	13-SEP-1999;	99US-0153758P.	PF	25-FEB-2000; 2000EP-00301439.
PR	15-SEP-1999;	99US-0154018P.	XX	
PR	16-SEP-1999;	99US-0154039P.	XX	
PR	20-SEP-1999;	99US-0154779P.	PR	25-FEB-1999;
PR	22-SEP-1999;	99US-0155139P.	PR	99US-0121825P.
PR	23-SEP-1999;	99US-0155486P.	PR	99US-0123180P.
PR	24-SEP-1999;	99US-0155659P.	PR	99US-0123548P.
PR	28-SEP-1999;	99US-0156458P.	PR	99US-0125788P.
PR	29-SEP-1999;	99US-0156596P.	PR	99US-0126264P.
PR	04-OCT-1999;	99US-0157117P.	PR	99US-0126785P.
PR	05-OCT-1999;	99US-0157753P.	PR	99US-0127462P.
PR	06-OCT-1999;	99US-0158653P.	PR	99US-0128234P.
PR	07-OCT-1999;	99US-0158029P.	PR	99US-0128714P.
PR	08-OCT-1999;	99US-0158232P.	PR	99US-0129845P.
PR	12-OCT-1999;	99US-0158369P.	PR	99US-0130077P.
PR	13-OCT-1999;	99US-0159293P.	PR	99US-0130449P.
PR	13-OCT-1999;	99US-0159294P.	PR	99US-0130510P.
PR	13-OCT-1999;	99US-0159295P.	PR	99US-0130891P.
PR	14-OCT-1999;	99US-0159329P.	PR	99US-0131449P.
PR	14-OCT-1999;	99US-0159330P.	PR	99US-0132048P.
PR	14-OCT-1999;	99US-0159331P.	PR	99US-0132407P.
PR	14-OCT-1999;	99US-0159637P.	PR	99US-0132484P.
PR	16-OCT-1999;	99US-0159584P.	PR	99US-0132485P.
PR	21-OCT-1999;	99US-0160741P.	PR	99US-0132486P.
PR	21-OCT-1999;	99US-0160767P.	PR	99US-0132487P.
PR	21-OCT-1999;	99US-0160768P.	PR	99US-0132863P.
PR	21-OCT-1999;	99US-0160770P.	PR	99US-0134256P.
PR	21-OCT-1999;	99US-0160814P.	PR	99US-0134218P.
PR	21-OCT-1999;	99US-0160815P.	PR	99US-0134219P.
PR	22-OCT-1999;	99US-0160980P.	PR	99US-0134219P.
PR	22-OCT-1999;	99US-0160981P.	PR	99US-0134370P.
PR	22-OCT-1999;	99US-0160989P.	PR	99US-0134768P.
PR	25-OCT-1999;	99US-0161404P.	PR	99US-0134941P.
PR	25-OCT-1999;	99US-0161405P.	PR	99US-0135124P.
PR	25-OCT-1999;	99US-0161406P.	PR	99US-0135353P.
PR	26-OCT-1999;	99US-0161359P.	PR	99US-0135629P.
PR	26-OCT-1999;	99US-0161360P.	PR	99US-0136021P.
PR	26-OCT-1999;	99US-0161361P.	PR	99US-0136392P.
PR	28-OCT-1999;	99US-0161820P.	PR	99US-0136782P.
PR	28-OCT-1999;	99US-0161922P.	PR	99US-0137222P.
PR	28-OCT-1999;	99US-0161993P.	PR	99US-0137528P.
PR	29-OCT-1999;	99US-0162142P.	PR	99US-0137502P.
PR			PR	99US-0137724P.
PR			PR	99US-0138094P.
PR			PR	99US-0138540P.
PR			PR	99US-0138847P.
PR			PR	99US-0139119P.
PR			PR	99US-0139452P.
PR			PR	99US-0139453P.
PR			PR	99US-0139492P.
PR			PR	99US-0139454P.
PR			PR	99US-0139455P.
PR			PR	99US-0139456P.
PR			PR	99US-0139457P.
PR			PR	99US-0139458P.
PR			PR	99US-0139459P.
PR			PR	99US-0139460P.
PR			PR	99US-0139461P.
PR			PR	99US-0139462P.
PR			PR	99US-0139463P.
PR			PR	99US-0139750P.
PR			PR	99US-0139763P.
PR			PR	99US-0139817P.
PR			PR	99US-0139899P.
PR			PR	99US-0140353P.
PR			PR	99US-0140354P.
PR			PR	99US-0140695P.
PR			PR	99US-0140823P.
PR			PR	99US-0140991P.
PR			PR	99US-0141287P.
PR			PR	99US-0141842P.

Query Match 50.8%; Score 135; DB 3; Length 426;
Best Local Similarity 46.9%; Pred. No. 1.2e-06;
Matches 23; Conservative 5; Mismatches 15; Indels 6; Gaps 1;
QY 2 CIIQDILLHCVSLQPCMKHTFCACVSGWHERSS-----LCPTCRCPV 44
Db 98 CSICLNHWDVTAAPCLHNFNGCFSEWNRSEKHKHVLCPQCRVV 146

RESULT 14
AAG20352
ID AAG20352 standard; protein; 453 AA.
XX
AC AAG20352;
XX
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 22505.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
CS Arabidopsis thaliana.
XX

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127482P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130831P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 18-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.

PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	21-JUN-1999;	99US-0139899P.
PR	22-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	23-JUN-1999;	99US-0140695P.
PR	24-JUN-1999;	99US-0140823P.
PR	28-JUN-1999;	99US-0140991P.
PR	28-JUN-1999;	99US-0141287P.
PR	30-JUN-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142254P.
PR	01-JUL-1999;	99US-0142255P.
PR	02-JUL-1999;	99US-0142390P.
PR	06-JUL-1999;	99US-0142803P.
PR	08-JUL-1999;	99US-0142920P.
PR	09-JUL-1999;	99US-0142977P.
PR	12-JUL-1999;	99US-0143542P.
PR	13-JUL-1999;	99US-0143624P.
PR	14-JUL-1999;	99US-0144005P.
PR	15-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	16-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-01443352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	23-JUL-1999;	99US-0145276P.
PR	26-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	02-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	14-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161408P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
PR	29-OCT-1999;	99US-0162143P.
Query Match 50.8%; Score 135; DB 3; Length 426;		
Best Local Similarity 46.9%; Pred. No. 1.2e-06;		
Matches 23; Conservative 5; Mismatches 15; Indels 6; Gaps 1;		
QY	2	CIICODLHDCVSLQPCVHTFCAACVSGWMERS-----LCPTCRCPV 44
DB	98	CSICLNHEDVVTAAFLNFCNCFSEWMRSEEEKHKHVLCPQCRITV 146
RESULT 13		
AAG37659		
ID	AAG37659 standard; protein; 426 AA.	
XX	AAG37659;	
AC	(first entry)	
XX	18-OCT-2000	
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 46343.	

PR	16-AUG-1999;	99US-0149368P.	ID	XX	AAG20353 standard; protein; 426 AA.
PR	17-AUG-1999;	99US-0149175P.	XX	AC	AAG20353;
PR	18-AUG-1999;	99US-0149426P.	XX	XX	17-OCT-2000 (first entry)
PR	20-AUG-1999;	99US-0149722P.	DT	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 22506.
PR	20-AUG-1999;	99US-0149723P.	XX	DE	Protein identification; signal transduction pathway; metabolic pathway;
PR	23-AUG-1999;	99US-0149929P.	XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	23-AUG-1999;	99US-0149930P.	XX	KW	termination sequence.
PR	25-AUG-1999;	99US-0150566P.	XX	OS	Arabidopsis thaliana.
PR	26-AUG-1999;	99US-0150884P.	XX	PN	EP1033405-A2.
PR	27-AUG-1999;	99US-0151065P.	XX	XX	06-SEP-2000.
PR	27-AUG-1999;	99US-0151066P.	XX	XX	25-FEB-2000; 2000EP-00301439.
PR	27-AUG-1999;	99US-0151080P.	XX	XX	25-FEB-1999; 99US-0121825P.
PR	30-AUG-1999;	99US-0151303P.	PR	PR	05-MAR-1999; 99US-0123180P.
PR	31-AUG-1999;	99US-0151438P.	PR	PR	05-MAR-1999; 99US-0123548P.
PR	01-SEP-1999;	99US-0151930P.	PR	PR	23-MAR-1999; 99US-0125788P.
PR	07-SEP-1999;	99US-0152363P.	PR	PR	23-MAR-1999; 99US-0126264P.
PR	10-SEP-1999;	99US-0153070P.	PR	PR	23-MAR-1999; 99US-0126785P.
PR	13-SEP-1999;	99US-0153758P.	PR	PR	01-APR-1999; 99US-0127462P.
PR	15-SEP-1999;	99US-0154018P.	PR	PR	06-APR-1999; 99US-0128234P.
PR	16-SEP-1999;	99US-0154039P.	PR	PR	08-APR-1999; 99US-0128714P.
PR	22-SEP-1999;	99US-0154779P.	PR	PR	16-APR-1999; 99US-0129845P.
PR	22-SEP-1999;	99US-0155139P.	PR	PR	19-APR-1999; 99US-0130077P.
PR	24-SEP-1999;	99US-0155659P.	PR	PR	21-APR-1999; 99US-0130449P.
PR	28-SEP-1999;	99US-0156458P.	PR	PR	23-APR-1999; 99US-0130510P.
PR	29-SEP-1999;	99US-0156596P.	PR	PR	23-APR-1999; 99US-0130891P.
PR	04-OCT-1999;	99US-0157117P.	PR	PR	28-APR-1999; 99US-0131449P.
PR	05-OCT-1999;	99US-0157753P.	PR	PR	30-APR-1999; 99US-0132048P.
PR	06-OCT-1999;	99US-0157865P.	PR	PR	04-MAY-1999; 99US-0132407P.
PR	07-OCT-1999;	99US-0158029P.	PR	PR	05-MAY-1999; 99US-0132484P.
PR	08-OCT-1999;	99US-0158232P.	PR	PR	06-MAY-1999; 99US-0132486P.
PR	12-OCT-1999;	99US-0158369P.	PR	PR	06-MAY-1999; 99US-0132487P.
PR	13-OCT-1999;	99US-0159293P.	PR	PR	07-MAY-1999; 99US-0132863P.
PR	13-OCT-1999;	99US-0159295P.	PR	PR	11-MAY-1999; 99US-0134256P.
PR	14-OCT-1999;	99US-0159322P.	PR	PR	14-MAY-1999; 99US-0134218P.
PR	14-OCT-1999;	99US-0159331P.	PR	PR	14-MAY-1999; 99US-0134219P.
PR	14-OCT-1999;	99US-0159337P.	PR	PR	14-MAY-1999; 99US-0134221P.
PR	14-OCT-1999;	99US-0159638P.	PR	PR	14-MAY-1999; 99US-0134370P.
PR	18-OCT-1999;	99US-0159584P.	PR	PR	18-MAY-1999; 99US-0134768P.
PR	21-OCT-1999;	99US-0160741P.	PR	PR	19-MAY-1999; 99US-0134941P.
PR	21-OCT-1999;	99US-0160767P.	PR	PR	20-MAY-1999; 99US-0135124P.
PR	21-OCT-1999;	99US-0160768P.	PR	PR	21-MAY-1999; 99US-0135353P.
PR	21-OCT-1999;	99US-0160770P.	PR	PR	24-MAY-1999; 99US-0135629P.
PR	21-OCT-1999;	99US-0160814P.	PR	PR	25-MAY-1999; 99US-0136021P.
PR	21-OCT-1999;	99US-0160815P.	PR	PR	27-MAY-1999; 99US-0136392P.
PR	22-OCT-1999;	99US-0160980P.	PR	PR	28-MAY-1999; 99US-0136782P.
PR	22-OCT-1999;	99US-0160981P.	PR	PR	01-JUN-1999; 99US-0137222P.
PR	22-OCT-1999;	99US-0160983P.	PR	PR	03-JUN-1999; 99US-0137528P.
PR	25-OCT-1999;	99US-0161404P.	PR	PR	04-JUN-1999; 99US-0137502P.
PR	25-OCT-1999;	99US-0161405P.	PR	PR	07-JUN-1999; 99US-0137724P.
PR	25-OCT-1999;	99US-0161406P.	PR	PR	08-JUN-1999; 99US-0138094P.
PR	26-OCT-1999;	99US-0161359P.	PR	PR	10-JUN-1999; 99US-0138540P.
PR	26-OCT-1999;	99US-0161360P.	PR	PR	10-JUN-1999; 99US-0138847P.
PR	26-OCT-1999;	99US-0161361P.	PR	PR	14-JUN-1999; 99US-0139119P.
PR	26-OCT-1999;	99US-0161362P.	PR	PR	16-JUN-1999; 99US-0139452P.
PR	28-OCT-1999;	99US-0161920P.	PR	PR	16-JUN-1999; 99US-0139453P.
PR	28-OCT-1999;	99US-0161992P.	PR	PR	17-JUN-1999; 99US-0139492P.
PR	28-OCT-1999;	99US-0161993P.	PR	PR	18-JUN-1999; 99US-0139454P.
PR	29-OCT-1999;	99US-0162142P.	PR	PR	18-JUN-1999; 99US-0139455P.
PR	29-OCT-1999;	99US-0162143P.	PR	PR	18-JUN-1999; 99US-0139456P.
PR	29-OCT-1999;	99US-0162144P.	PR	PR	18-JUN-1999; 99US-0139457P.
PR	29-OCT-1999;	99US-0162145P.	PR	PR	18-JUN-1999; 99US-0139458P.
PR	29-OCT-1999;	99US-0162146P.	PR	PR	18-JUN-1999; 99US-0139459P.
PR	29-OCT-1999;	99US-0162147P.	PR	PR	18-JUN-1999; 99US-0139460P.

Query Match 50.8%; Score 135; DB 3; Length 350;
Best Local Similarity 46.9%; Pred. No. 9.9e-07;
Matches 23; Conservative 5; Mismatches 15; Indels 6; Gaps 1;
QY 2 CIIICQLHDCVSLQPCMHITPCAACTGHWERS-----LCPTCRCPV 44
DB 22 CSICLINIWDVVTAAFLNFCNGCFSEWNRSEKHKHVLCPQCRITV 70

PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142302P.
PR 08-JUL-1999; 99US-0142803P.
PR 08-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 13-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 08-OCT-1999; 99US-0157855P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161952P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match : 50.8%; Score 135; DB 3; Length 350;
Best Local Similarity 46.9%; Pred. No. 9.9e-07;
Matches 23; Conservative 5; Mismatches 15; Indels 6; Gaps 1;

26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
22-AUG-2000; 2000US-0226688P.
22-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
05-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
08-SEP-2000; 2000US-0229513P.
08-SEP-2000; 2000US-0231413P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
27-SEP-2000; 2000US-0235834P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239335P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0244617P.
12-NOV-2000; 2000US-0249239P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;
WPI; 2003-147444/14.
N-PSDB; ABX73645.

New polypeptides and nucleic acids, useful in gene therapy for treating, inhibiting or preventing e.g. neural, immune system, muscular, respiratory, reproductive, Gastrointestinal, pulmonary, cardiovascular or renal disorders.

Claim 11; SEQ ID NO 1270; 402pp; English.

The invention relates to human novel polypeptides and their associated polynucleotides. The polypeptides and polynucleotides are useful in gene therapy for treating, inhibiting or preventing neural disorders, immune system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis), muscular disorders, respiratory diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders, gastrointestinal disorders, pulmonary disorders, cardiovascular disorders (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute kidney failure and end-stage renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and leukemia), inflammatory diseases (e.g. septic shock, bursitis and appendicitis), allergic reactions and conditions (e.g. asthma), blood related disorders (e.g. thrombosis, atherosclerosis and myocardial infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and ABU55748 represent human novel polypeptides of the invention

XX SQ Sequence 230 AA;
Query Match 66.2%; Score 176; DB 6; Length 230;
Best Local Similarity 100.0%; Pred. No. 2,7e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 PMETFCACYSGWMMERSLICPTCRCPV 44
DB 5 PMETFCACYSGWMMERSLICPTCRCPV 32

RESULT 10
AAG20354
ID AAG20354 standard; protein; 350 AA.
XX
AC AAG20354;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22507.
XX
KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; Genetic mapping; gene expression control; promoter; termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 09-MAR-1999; 99US-0123180P.
PR 05-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 06-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131443P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134258P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135623P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.

PT	used as food additives or preservatives.	PT	used as food additives or preservatives.
XX	Claim 11; SEQ ID NO 1270; 980pp; English.	XX	Claim 11; SEQ ID NO 1270; 980pp; English.
PS	The invention relates to isolated nucleic acid molecules and their	PS	The invention relates to isolated nucleic acid molecules and their
XX	encoded secreted proteins. The nucleic acids and proteins are used to	XX	encoded secreted proteins. The nucleic acids and proteins are used to
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,	CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used	CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC	in diagnosing a pathological condition or susceptibility to a	CC	in diagnosing a pathological condition or susceptibility to a
CC	pathological condition. Antibodies to the proteins can also be used in	CC	pathological condition. Antibodies to the proteins can also be used in
CC	alleviating symptoms associated with the disorders and in diagnostic	CC	alleviating symptoms associated with the disorders and in diagnostic
CC	immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays	CC	immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC	(ELISA). Disorders which are diagnosed or treated include autoimmune	CC	(ELISA). Disorders which are diagnosed or treated include autoimmune
CC	diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.	CC	diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC	neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac	CC	neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC	arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,	CC	arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC	nervous system disorders e.g. Alzheimer's disease, infections caused by	CC	nervous system disorders e.g. Alzheimer's disease, infections caused by
CC	bacteria, viruses and fungi and ocular disorders e.g. corneal infection,	CC	bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC	and many other disorders listed in the specification. The polypeptides	CC	and many other disorders listed in the specification. The polypeptides
CC	can also be used to aid wound healing and epithelial cell proliferation,	CC	can also be used to aid wound healing and epithelial cell proliferation,
CC	to prevent skin aging due to sunburn, to maintain organs before	CC	to prevent skin aging due to sunburn, to maintain organs before
CC	transplantation, for supporting cell culture of primary tissues, to	CC	transplantation, for supporting cell culture of primary tissues, to
CC	regenerate tissues and in chemotaxis. The polypeptides can also be used	CC	regenerate tissues and in chemotaxis. The polypeptides can also be used
CC	as a food additive or preservative to increase or decrease storage	CC	as a food additive or preservative to increase or decrease storage
CC	capabilities, fat content, lipid, protein, carbohydrate, vitamins,	CC	capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC	minerals, cofactors and other nutritional components. The present	CC	minerals, cofactors and other nutritional components. The present
CC	sequence represents a novel secreted protein of the invention. Note: The	CC	sequence represents a novel secreted protein of the invention. Note: The
CC	sequence data for this patent did not form part of the printed	CC	sequence data for this patent did not form part of the printed
Query Match 66.2%; Score 176; DB 4; Length 230;			
Best Local Similarity 100.0%; Pred. No. 2.7e-11;			
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	17 PCWHFTCAACYSYGMWERSLCTCRCPV 44	QY	17 PCWHFTCAACYSYGMWERSLCTCRCPV 44
Db	5 PCWHFTCAACYSYGMWERSLCTCRCPV 32	Db	5 PCWHFTCAACYSYGMWERSLCTCRCPV 32
RESULT 9			
ABU55386	ID ABU55386 standard; protein; 230 AA.	ABU55386	ID ABU55386 standard; protein; 230 AA.
XX	ABU55386;	XX	ABU55386;
AC	18-MAR-2003 (first entry)	AC	18-MAR-2003 (first entry)
XX	Human novel polypeptide #473.	XX	Human novel polypeptide #473.
DE	Human; neural disorder; immune system disorder; renal disorder;	DE	Human; neural disorder; immune system disorder; renal disorder;
XX	muscular disorder; respiratory disease; reproductive disorder;	XX	muscular disorder; respiratory disease; reproductive disorder;
KW	gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;	KW	gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW	hyperproliferative disorder; inflammatory disease; allergic reaction;	KW	hyperproliferative disorder; inflammatory disease; allergic reaction;
KW	blood related disorder; cancer; immunosuppressive; antiinflammatory;	KW	blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW	cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;	KW	cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW	haemostatic; antiarteriosclerotic.	KW	haemostatic; antiarteriosclerotic.
XX	Homo sapiens.	XX	Homo sapiens.
OS	US20021132753-A1.	OS	US20021132753-A1.
XX	19-SEP-2002.	XX	19-SEP-2002.
PN	17-JAN-2001; 2001US-00764864.	PN	17-JAN-2001; 2001US-00764864.
PD	31-JAN-2000; 2000US-0179065P.	PD	31-JAN-2000; 2000US-0179065P.
XX	04-FEB-2000; 2000US-0180628P.	XX	04-FEB-2000; 2000US-0180628P.
PR	28-JUN-2000; 2000US-0214886P.	PR	28-JUN-2000; 2000US-0214886P.
PR	07-JUL-2000; 2000US-0216647P.	PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216800P.	PR	07-JUL-2000; 2000US-0216800P.
PR	11-JUL-2000; 2000US-0217487P.	PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217496P.	PR	11-JUL-2000; 2000US-0217496P.
PR	14-JUL-2000; 2000US-0218290P.	PR	14-JUL-2000; 2000US-0218290P.
PR	26-JUL-2000; 2000US-0220963P.	PR	26-JUL-2000; 2000US-0220963P.
PA	(HUMA-) HUMAN GENOME SCI INC.	PA	(HUMA-) HUMAN GENOME SCI INC.
XX	Rosen CA, Barash SC, Ruben SM;	XX	Rosen CA, Barash SC, Ruben SM;
PI	WPI; 2001-488783/53.	PI	WPI; 2001-488783/53.
XX	N-PSDE; AAS26304.	XX	N-PSDE; AAS26304.
DR	New nucleic acid molecules encoding 461 human secreted proteins for	DR	New nucleic acid molecules encoding 461 human secreted proteins for
XX	diagnosing, preventing, treating or ameliorating medical conditions and	XX	diagnosing, preventing, treating or ameliorating medical conditions and
PT		PT	

WPI; 2001-182927/18.
N-PSDB; AAF30352.

Novel nucleic acid sequence of mitotic checkpoint gene encoding a checkpoint with forkhead-associated domain and ring finger protein, for diagnosing tumorigenic cells and in screening for anticancer drugs.

Claim 8(a); Fig 4A-C; 85pp; English.

The present sequence is that of human mitotic checkpoint protein Chfr, having a forkhead associated domain (FHA) and a ring finger domain. The protein is required for regulation of the transition of cells from prophase to metaphase during mitosis. It has ubiquitin-protein ligase activity. The Chfr checkpoint was evident in primary human cells, but was inactivated in 4 of 8 human cancer cell lines. In U2OS cells, a mutation was identified that caused a Val to Met amino acid substitution in the highly conserved C-terminal Cys-rich region of the Chfr protein. In the absence of the Chfr checkpoint, cells subjected to mitotic stress condensed their chromosomes despite failing to separate their chromosomes. Chfr may monitor centrosome separation. Inactivation of the chfr gene (see AAF30352) in human cancer is theorized to underlie the increased sensitivity of cancer cells to antimitotic drugs. Polypeptides comprising the present sequence, or sequences comprising at least amino acids 31-103, 303-346 and/or 476-641 of this sequence, are claimed. Claimed methods of determining the tumorigenic potential of a cell comprise examining the cell for the presence of Chfr expression or for the presence of Chfr-mediated ubiquitin-protein ligase activity (in both cases, absence of expression indicating predisposition to tumorigenesis upon exposure to mitotic stress). A diagnostic kit for detecting the tumorigenic potential of cell cells comprises may comprise a ligand that binds to Chfr, such as an antibody or its fragment. Inhibitors of Chfr are identified by monitoring their effect on Chfr expression, and are used to retard the growth of cancer cells

Sequence 664 AA;

Query Match 100.0%; Score 266; DB 4; Length 664;
Best Local Similarity 100.0%; Pred. No. 1.3e-20;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0

1 TCIIICDILHDCVSLQPCWHFFCAACYSGWNERSLQPTCRCPV 44
|||
303 TCIIICDILHDCVSLQPCWHFFCAACYSGWNERSLQPTCRCPV 346

RESULT 8
AAU16317
ID AAU16317 standard; protein; 230 AA.
XX AAU16317;
AC
XX
07-NOV-2001 (first entry)
XX
XX Human novel secreted protein, Seq ID 1270.
XX
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
XX preservative; antiproliferative.
XX Homo sapiens.
OS
XX
XX WO200155322-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
XX
XX

29-JUL-1999; 99JP-00248036.
 27-AUG-1999; 99JP-00300253.
 11-JAN-2000; 2000JP-00118776.
 02-MAY-2000; 2000JP-00183767.
 09-JUN-2000; 2000JP-00241899.
 (HELI-) HELIX RES INST.
 Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Iehii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 WPI; 2001-318749/34.
 Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.
 The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 XX Sequence 652 AA;
 Query Match 100.0%; Score 266; DB 4; Length 652;
 Best Local Similarity 100.0%; Pred. No. 1.3e-20;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TCIIQDLHDCVSLQPCWHTFCAACYSYSGWMSLSLCTCRCPV 44
 Db 291 TCIIQDLHDCVSLQPCWHTFCAACYSYSGWMSLSLCTCRCPV 334
 RESULT 6
 AAB3843
 ID AAB3843 standard; protein; 564 AA.
 XX AAB3843;
 AC AAB3843;
 XX 22-AUG-2001 (first entry)
 DT Amino acid sequence of a human ring finger protein designated FHAR1.
 DE FHAR1; RING finger protein; cancer; vaccine.
 KW Homo sapiens.
 OS
 XX WO200142430-A1.
 PN 14-JUN-2001.
 XX 07-DEC-2000; 2000WO-US033094.
 PF

XX 08-DEC-1999; 99US-00456876.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA Zhou B, Zhu Y, Chaturvedi P, Hurler MR, Li X;
 PI WPI; 2001-381663/40.
 DR N-PSDB; AAF89709.
 XX New FHAR1 polypeptide, a member of the RING finger protein family for diagnosing and treating cancer, and for use in anti-cancer vaccines.
 PT Claim 1; Page 19; 28pp; English.
 PS The present sequence represents a FHAR1 polypeptide, which is a member of the RING finger protein family. FHAR1 is useful in the treatment of cancer, and as a vaccine for inducing an immunological response in a mammal. FHAR1 polynucleotides may also be used as a diagnostic reagent through detection of mutations in the associated gene, and for chromosome localization studies, and tissue expression studies. FHAR1 antibodies are useful to isolate and to identify clones expressing the polypeptides, or to purify the polypeptides by affinity chromatography and to treat cancer
 XX Sequence 664 AA;
 SQ Query Match 100.0%; Score 266; DB 4; Length 664;
 Best Local Similarity 100.0%; Pred. No. 1.3e-20;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TCIIQDLHDCVSLQPCWHTFCAACYSYSGWMSLSLCTCRCPV 44
 Db 303 TCIIQDLHDCVSLQPCWHTFCAACYSYSGWMSLSLCTCRCPV 346
 RESULT 7
 AAB20219
 ID AAB20219 standard; protein; 664 AA.
 XX AAB20219;
 AC AAB20219;
 XX 14-MAY-2001 (first entry)
 DT Human Chfr (checkpoint with FHA and ring finger) protein.
 DE Checkpoint with forkhead associated domain and ring finger; Chfr; human; mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening; ubiquitin-protein ligase.
 KW Homo sapiens.
 OS
 XX Location/Qualifiers
 FH Domain 31..103
 FT /label= Forkhead-associated_domain
 FT Domain 303..346
 FT /label= Ring_finger-domain
 FT Region 476..641
 FT /note= "cysteine-rich region"
 FT Misc-difference 580
 FT /note= "Met in U2OS cells"
 XX WO200109150-A2.
 PN 08-FEB-2001.
 PD 14-JUN-2000; 2000WO-US016391.
 XX 29-JUL-1999; 99US-0146194P.
 PR (WIST-) WISTAR INST ANATOMY & BIOLOGY.
 PA Halazonetis T, Scolnick D;
 PI

DE Human protein sequence SEQ ID NO:12128.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
KW Homo sapiens.
OS Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248035.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX Claim 8; SEQ ID NO 12128; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92445 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX Sequence 623 AA;
XX Query Match 100.0%; Score 266; DB 4; Length 623;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-20;
XX Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCIIQDILLHDCVSLQPCMTTFCACYSGWMSRSSLCPTCRCPV 44
DB 262 TCIIQDILLHDCVSLQPCMTTFCACYSGWMSRSSLCPTCRCPV 305
RESULT 4
ABB97233
ID ABB97233 standard; protein; 623 AA.
XX ABB97233;
AC ABB97233;
XX 27-JUN-2002 (first entry)
DT

XX Novel human protein SEQ ID NO: 501.
DE Human; antinaemic; vulnary; antiinflammatory; immunomodulator;
KW antinfertility; cerebroprotective; cyostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.
XX Homo sapiens.
XX WO200222560-A2.
XX 21-MAR-2002.
XX 10-SEP-2001; 2001WO-US026015.
XX 11-SEP-2000; 2000US-00659671.
XX (HYSE-) HYSEQ INC.
XX Tang YN, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmahac RT;
XX WPI; 2002-292408/33.
XX N-PSDB; AEN32419.
XX An isolated polynucleotide for treating diseases associated with its encoded polypeptide such as cancer and multiple sclerosis.
XX Example 2; SEQ ID NO 501; 509pp; English.
XX The present invention provides the protein and coding sequences of 444 novel human proteins. These were isolated from expressed sequences tags (ESTs). They can be used to stimulate cell growth, to regulate haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth e.g. in burn treatment, to regulate the immune system e.g. to treat multiple sclerosis, to regulate activin or inhibin e.g. to treat infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke and cancer, to screen for drugs, to treat inflammatory conditions e.g. rheumatoid arthritis, and to treat nervous system disorders e.g. Parkinson's disease. The present sequence is a protein of the invention
XX Sequence 623 AA;
XX Query Match 100.0%; Score 266; DB 5; Length 623;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-20;
XX Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCIIQDILLHDCVSLQPCMTTFCACYSGWMSRSSLCPTCRCPV 44
DB 262 TCIIQDILLHDCVSLQPCMTTFCACYSGWMSRSSLCPTCRCPV 305
RESULT 5
AAB93168
ID AAB93168 standard; protein; 652 AA.
XX AAB93168;
AC AAB93168;
XX 26-JUN-2001 (first entry)
DT Human protein sequence SEQ ID NO:12100.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX

CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 100.0%; Score 266; DB 4; Length 426;
Best Local Similarity 100.0%; Pred. No. 9.2e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCIICQDLHDCVSLQPCMHFTCAACYSWGMWSSSLCPTCRCPV 44
DB 213 TCIICQDLHDCVSLQPCMHFTCAACYSWGMWSSSLCPTCRCPV 256

RESULT 2
ABU54925
ID ABU54925 standard; protein; 426 AA.

AC ABU54925;

DT 18-MAR-2003 (first entry)

DE Human novel polypeptide #12.

XX Human; neural disorder; immune system disorder; renal disorder;
XX muscular disorder; respiratory disease; reproductive disorder;
XX gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX hyperproliferative disorder; inflammatory disease; allergic reaction;
XX blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX cardiovascular; nephrotropic; cycostatic; antiallergic; thrombolytic;
XX haemostatic; antiarteriosclerotic.

XX Homo sapiens.

XX US2002132753-A1.

XX 19-SEP-2002.

XX 17-JAN-2001; 2001US-00764864.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 28-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0216847P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 22-AUG-2000; 2000US-0226868P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 01-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 05-SEP-2000; 2000US-0229513P.

XX 08-SEP-2000; 2000US-0231413P.

XX 21-SEP-2000; 2000US-0234223P.

XX 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 01-NOV-2000; 2000US-0241809P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX MPI; 2003-147444/14.
XX N-PSDB; ABX73184.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,
XX inhibiting or preventing e.g. neural, immune system, muscular,
XX respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
XX renal disorders.

XX Claim 11; SEQ ID NO 809; 402pp; English.

XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
XX (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
XX leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
XX appendicitis), allergic reactions and conditions (e.g. asthma), blood
XX related disorders (e.g. thrombosis, atherosclerosis and myocardial
XX infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
XX ABU55748 represent human novel polypeptides of the invention

XX Sequence 426 AA;

Query Match 100.0%; Score 266; DB 6; Length 426;
Best Local Similarity 100.0%; Pred. No. 9.2e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCIICQDLHDCVSLQPCMHFTCAACYSWGMWSSSLCPTCRCPV 44
DB 213 TCIICQDLHDCVSLQPCMHFTCAACYSWGMWSSSLCPTCRCPV 256

RESULT 3

AAB93182

ID AAB93182 standard; protein; 623 AA.

XX AAB93182;

XX 26-JUN-2001 (first entry)

XX

PR 14-AUG-2000; 2000US-02252113P.
PR 14-AUG-2000; 2000US-02252114P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 22-AUG-2000; 2000US-0227009P.
PR 30-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229511P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233397P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.

XX PA
XX PI
XX XX
DR DR
XX XX
PT PT
PT PT
XX XX
PS PS
XX XX
CC CC
CC CC
CC CC
CC CC
CC CC
CC CC
CC CC
CC CC
CC CC
CC CC
CC CC
CC CC

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-488783/53.
N-PSDB; AAS25843.
New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.
Claim 11; SEQ ID NO 809; 980pp; English.
The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radiimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infectious caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:36:57 ; Search time 9.01373 Seconds
(without alignments)
1379.240 Million cell updates/sec

Title: US-10-048-046-2_COPY_303_346

Perfect score: 266
Sequence: 1 TCICQDLHDCVSLQPCMH.....ACVSGWNERSLQPTCRCPV 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	266	100.0	426	AAU15856	AAU15856 Human nov
2	266	100.0	426	ABU54925	ABU54925 Human nov
3	266	100.0	623	ABU93182	ABU93182 Human pro
4	266	100.0	623	ABB97233	ABB97233 Novel hum
5	266	100.0	652	AAU93168	AAU93168 Human pro
6	266	100.0	664	AAU83843	AAU83843 Amino aci
7	266	100.0	664	AAU20219	AAU20219 Human Chf
8	176	66.2	230	AAU16317	AAU16317 Human nov
9	176	66.2	230	ABU53386	ABU53386 Human nov
10	135	50.8	350	AAU20354	AAU20354 Arabidops
11	135	50.8	350	AAU37660	AAU37660 Arabidops
12	135	50.8	426	AAU20353	AAU20353 Arabidops
13	135	50.8	426	AAU37659	AAU37659 Arabidops
14	135	50.8	453	AAU20352	AAU20352 Arabidops
15	135	50.8	453	AAU37658	AAU37658 Arabidops
16	129	48.5	67	AAU60624	AAU60624 Varicella
17	115	43.2	242	AAU39926	AAU39926 Human pol
18	115	43.2	247	AAU77120	AAU77120 Human pro
19	115	43.2	254	ABU79353	ABU79353 Human ova
20	115	43.2	261	AAU41712	AAU41712 Human pol
21	113.5	42.7	321	AAU06408	AAU06408 Arabidops
22	113.5	42.7	322	AAU06407	AAU06407 Arabidops
23	113.5	42.7	426	AAU06406	AAU06406 Arabidops
24	112	42.1	242	AAU79067	AAU79067 Human pro
25	111	41.7	495	ABU56805	ABU56805 Human sec

26	110	41.4	498	2	AAW80439	Feline he
27	108.5	40.8	485	2	AAU95017	Human pro
28	107	40.2	551	2	AAW27150	HMG-CoA r
29	104	39.1	330	4	AAE01417	Human sec
30	103	38.7	328	4	AAE01418	Human alb
31	103	38.7	489	5	ABG64061	Human alb
32	103	38.7	505	4	AAE01380	Human gen
33	103	38.7	616	5	ABU83078	Synovioli
34	103	38.7	617	4	AAE01355	Human gen
35	103	38.7	617	5	ABG79641	Human nov
36	103	38.7	617	5	ABG64063	Human alb
37	103	38.7	617	5	ABU83074	Synovioli
38	103	38.7	617	6	AAE37275	Human gen
39	102	38.3	328	4	ABU64164	Drosophi
40	101	38.0	568	3	AAU18230	Plasmodiu
41	100.5	37.8	499	5	AAO22598	499-mer S
42	100.5	37.8	1208	4	AAU52592	Human pro
43	100.5	37.8	1213	5	AAO22597	1213-mer
44	100	37.6	222	4	ABU60472	Drosophi
45	100	37.6	356	6	ABU65227	Novel hum

ALIGNMENTS

RESULT 1
AAU15856
ID AAU15856 standard; protein; 426 AA.
XX AC AAU15856;
XX DT 07-NOV-2001 (first entry)
XX DE Human novel secreted protein, Seq ID 809.
XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX OS Homo sapiens.
XX WO200155322-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001341.
XX PP 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.

MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
RA Villalón D.K., Wuzly D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywiński M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RA and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RA [3]
RA SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RA EMBL; AF076687; AAD20943.1; -
RA EMBL; BC055061; AAH5061.1; -
RA MGD; MGI:1351625; Nbn.
RA GO; GO:0005634; C:nucleus; IDA.
RA InterPro; IPR001357; BRCT.
RA InterPro; IPR000253; FHA.
RA InterPro; IPR008984; SMAD_FHA.
RA Pfam; PF00533; BRCT; 1.
RA Pfam; PF00498; FHA; 1.
RA SMART; SM00292; BRCT; 1.
RA SMART; SM00240; FHA; 1
RA PROSITE; PS50006; FHA_DOMAIN; 1.
RA Hypothetical protein.
RA SEQUENCE 751 AA; 83794 MW; C9F597CC08227B2C CRC64;
RA KW

Query Match 23.0%; Score 87; DB 11; Length 751;
Best Local Similarity 26.6%; Pred. No. 0.15;
Matches 21; Conservative 21; Mismatches 27; Indels 10; Gaps 2;

QY 2 LLRKREWTGRRRGCDLSFSPSNKLVSGDHCRIIVD-----EKSGQVTLTSTSGT 52

pb 18 LLAGEVIVVG-RKNCGLIENDQSIENHAVLTWFFVTSLSQTDIPIPTIKNSKYGT 76

Qy 53 VINKLVVKKQTCP LQTGD 71
: : : : :
Db 77 FVNEEKMQTGLSCTLKTGD 95

RESULT 15

ID	Q8G6Q4	PRELIMINARY;	PRT;	176 AA.
AC	Q8G6Q4;			
DC	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Hypothetical protein with FHA domain.			
GN	El0584			
OS	Bifidobacterium longum.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;			
OC	Bifidobacteriaceae; Bifidobacterium.			
ON	NCBI_TaxID=216916;			
OX	[1]			
PC	SEQUENCE FROM N. A.			
RP	STRAIN=NCC 2705;			
RX	MEDLINE=22284977; PubMed=12381787;			
RA	Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,			
RA	Fessi G., Zwahten M.-C., Desiere F., Bork P., Delley M.,			

RA	Pridmore R.D., Arigoni F.;
RT	"The genome sequence of Fibrobacterium longum reflects its adaptation
RT	to the human gastrointestinal tract.",
RL	Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR	EMBL; AB014680; AANZ4408.1; -.
DR	InterPro; IPR000253; FHA.
DR	InterPro; IPR008984; SWAD_FHA.
DR	Pfam; PF00498; FHA; 1.
DR	SMART; SM00240; FHA; 1.
DR	PROSITE; PS50006; FHA_DOMAIN; 1.
DR	Hypothetical protein; Complete proteome.
KW	SEQUENCE 176 AA; 19530 MW; 1612CDF8C9A93CF6 CRC64;
SQ	
Query Match	22.8%; Score 86; DB 16; Length 176;
Best Local Similarity	32.8%; Pred. No. 0.035;
Matches	22; Conservative 15; Mismatches 28; Indels 2; Gaps 2;
QY	1 VLLKRWNTGRRRGDLSFPSSNKLVSGBHCHIVVDEKSQGVTLSD-TSTSGTWINKLV 59 : : : : : : : : : : : : : : : : : : : : : : : : 96 VPLAEADITLGRAASNTVMVL-DDEFVSSHARVIYRDSQWAIEDLNSTNGTVVNQORI 154
Db	60 VKKQTCP 66
QY	:
Db	155 NRPTLP 161

Search completed: May 7, 2004, 14:49:02
Job time : 11.5607 secs

Query Match 23.0%; Score 87; DB 11; Length 751;

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain stem;
 RA Mas C., Bourgeois F., Simonneau M.;
 RT "Isolation of 50 cDNAs differentially expressed in embryonic forebrain
 as compared to mid and hindbrain: a strategy to identify candidate
 genes involved in human neurodevelopmental diseases.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF092840; AAC62113.1; -.
 DR MGD; MGI:1351625; Nbn.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR001357; BRCT.
 DR InterPro; IPR000253; FHA.
 DR InterPro; IPR008984; SNAD_FHA.
 DR Pfam; PF00533; BRCT; 1.
 DR Pfam; PF00498; BRCT; 1.
 DR SMART; SM00292; BRCT; 1.
 DR SMART; SM00240; FHA; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 SQ SEQUENCE 751 AA; 83826 MW; 9D136B8C23DC51D9 CRC64;
 Query Match 23.0%; Score 87; DB 11; Length 751;
 Best Local Similarity 26.6%; Pred. No. 0.15;
 Matches 21; Conservative 21; Mismatches 27; Indels 10; Gaps 2;
 QY 2 LLRKKEWTIGRRGCDLSPFNKLVSGDHCRIYVD-----EKSGQVLEDTSTSGT 52
 DB 18 LLAGVEYVVG-RKNCGLIENDQSISRHAFLTVNFPVTSLSQTDIEPTLIKDNKYGT 76
 QY 53 VINKLVKVKQTCPLQTGD 71
 DB 77 FVNEEKMQTGLSCTLKTKGD 95
 RESULT 13
 Q81116 PRELIMINARY; PRT; 751 AA.
 ID Q81116;
 AC Q81116;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Nibrin.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC044773; AAH44773.1; -.
 DR InterPro; IPR000253; FHA.
 DR InterPro; IPR008984; SNAD_FHA.
 DR Pfam; PF00498; FHA; 1.
 DR SMART; SM00240; FHA; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 SQ SEQUENCE 751 AA; 83829 MW; DD78D11F3C1AE3F5 CRC64;
 Query Match 23.0%; Score 87; DB 11; Length 751;
 Best Local Similarity 26.6%; Pred. No. 0.15;
 Matches 21; Conservative 21; Mismatches 27; Indels 10; Gaps 2;
 QY 2 LLRKKEWTIGRRGCDLSPFNKLVSGDHCRIYVD-----EKSGQVLEDTSTSGT 52
 DB 18 LLAGVEYVVG-RKNCGLIENDQSISRHAFLTVNFPVTSLSQTDIEPTLIKDNKYGT 76
 QY 53 VINKLVKVKQTCPLQTGD 71
 DB 77 FVNEEKMQTGLSCTLKTKGD 95
 RESULT 14
 Q9R207 PRELIMINARY; PRT; 751 AA.
 ID Q9R207;
 AC Q9R207;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Nibrin (Hypothetical protein).
 GN NBN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20108791; PubMed=10640816;
 RA Vissinga C.S., Yeo T.C., Woessner J., Massa H.F., Wilson R.K.,
 RA Trask B.J., Concannon P.;
 RT "Identification, characterization, and mapping of a mouse homolog of
 the gene mutated in Nijmegen breakage syndrome.";
 RL Cytogenet. Cell Genet. 87:80-84(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP SEQUENCE=FVB/N; TISSUE=Colon;
 RC

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain stem;
 RA Mas C., Bourgeois F., Simonneau M.;
 RT "Isolation of 50 cDNAs differentially expressed in embryonic forebrain
 as compared to mid and hindbrain: a strategy to identify candidate
 genes involved in human neurodevelopmental diseases.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF092840; AAC62113.1; -.
 DR MGD; MGI:1351625; Nbn.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR InterPro; IPR001357; BRCT.
 DR InterPro; IPR000253; FHA.
 DR InterPro; IPR008984; SNAD_FHA.
 DR Pfam; PF00533; BRCT; 1.
 DR Pfam; PF00498; BRCT; 1.
 DR SMART; SM00292; BRCT; 1.
 DR SMART; SM00240; FHA; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 SQ SEQUENCE 751 AA; 83826 MW; 9D136B8C23DC51D9 CRC64;
 Query Match 23.0%; Score 87; DB 11; Length 751;
 Best Local Similarity 26.6%; Pred. No. 0.15;
 Matches 21; Conservative 21; Mismatches 27; Indels 10; Gaps 2;
 QY 2 LLRKKEWTIGRRGCDLSPFNKLVSGDHCRIYVD-----EKSGQVLEDTSTSGT 52
 DB 18 LLAGVEYVVG-RKNCGLIENDQSISRHAFLTVNFPVTSLSQTDIEPTLIKDNKYGT 76
 QY 53 VINKLVKVKQTCPLQTGD 71
 DB 77 FVNEEKMQTGLSCTLKTKGD 95
 RESULT 13
 Q81116 PRELIMINARY; PRT; 751 AA.
 ID Q81116;
 AC Q81116;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Nibrin.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC044773; AAH44773.1; -.
 DR InterPro; IPR000253; FHA.
 DR InterPro; IPR008984; SNAD_FHA.
 DR Pfam; PF00498; FHA; 1.
 DR SMART; SM00240; FHA; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 DR PROSITE; PS00006; FHA_DOMAIN; 1.
 SQ SEQUENCE 751 AA; 83829 MW; DD78D11F3C1AE3F5 CRC64;
 Query Match 23.0%; Score 87; DB 11; Length 751;
 Best Local Similarity 26.6%; Pred. No. 0.15;
 Matches 21; Conservative 21; Mismatches 27; Indels 10; Gaps 2;
 QY 2 LLRKKEWTIGRRGCDLSPFNKLVSGDHCRIYVD-----EKSGQVLEDTSTSGT 52
 DB 18 LLAGVEYVVG-RKNCGLIENDQSISRHAFLTVNFPVTSLSQTDIEPTLIKDNKYGT 76
 QY 53 VINKLVKVKQTCPLQTGD 71
 DB 77 FVNEEKMQTGLSCTLKTKGD 95
 RESULT 14
 Q9R207 PRELIMINARY; PRT; 751 AA.
 ID Q9R207;
 AC Q9R207;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Nibrin (Hypothetical protein).
 GN NBN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20108791; PubMed=10640816;
 RA Vissinga C.S., Yeo T.C., Woessner J., Massa H.F., Wilson R.K.,
 RA Trask B.J., Concannon P.;
 RT "Identification, characterization, and mapping of a mouse homolog of
 the gene mutated in Nijmegen breakage syndrome.";
 RL Cytogenet. Cell Genet. 87:80-84(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP SEQUENCE=FVB/N; TISSUE=Colon;
 RC

Q87PUO PRELIMINARY; PRT; 504 AA.
AC Q87PUO;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN VFL111.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:H6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RA "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
RL EMBL; AP005078; BAC59674.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 504 AA; 55469 MW; 2F96C75A4F83648E CRC64;

Query Match 28.4%; Score 107.5; DB 16; Length 504;
Best Local Similarity 36.4%; Pred. No. 0.00034;
Matches 24; Conservative 12; Mismatches 27; Indels 3; Gaps 2;

QY 9 TIGRRCGDLSPS-NKLVSCHCRIVVDEKSGQVLTSTSGTVINKLVKVKQTCPL 67
Db 31 SVGRASCTLSLTHRNFRISGTHCLISV--YGTFFISDVSTGTVWNGKILKNQIPSI 88

QY 68 QTGDVI 73
Db 89 VEGDVV 94

RESULT 9
Q97LQO PRELIMINARY; PRT; 159 AA.
AC Q97LQO;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FHA-domain containing secreted protein.
GN CAC0504.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Neelling J., Breton G., Onelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe B., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
RL EMBL; AE007565; AAK78484.1; -.
DR PIR; A96962; A96962.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.

KW Complete proteome.
SQ SEQUENCE 159 AA; 17667 MW; F4P44A3D3BF05AA1 CRC64;

Query Match 23.3%; Score 88; DB 16; Length 159;
Best Local Similarity 41.4%; Pred. No. 0.018;
Matches 29; Conservative 12; Mismatches 21; Indels 8; Gaps 5;

QY 6 REWTIGRRGCDLSPSNK-LVSGDHCRIVVDEKSGQVLTSTSGTVINKLVKVKQ 63
Db 85 REITIGRKD--DASIMLNEGVSCHHARVYL--RNNQYILEDLSNGTVLNQKI--KS 138

QY 64 TCPLQTCGVI 73
Db 139 KAVIKSGDEI 148

RESULT 10
Q8CCYO PRELIMINARY; PRT; 546 AA.
AC Q8CCYO;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Nibrin (Fragment).
GN NBN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RL EMBL; AK031933; BAC27610.1; -.
DR MGD; MGI:1351625; Nbn.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
FT NON_TER 546 546
SQ SEQUENCE 546 AA; 60269 MW; 69394A2537EPEF20 CRC64;

Query Match 23.0%; Score 87; DB 11; Length 546;
Best Local Similarity 26.6%; Pred. No. 0.1;
Matches 21; Conservative 21; Mismatches 27; Indels 10; Gaps 2;

QY 2 LLRKEWTIGRRGCDLSPSNKLVSGDHCRIVVD-----EKSGQVLTSTSGT 52
Db 18 LLAGVEYVVG-RKNGILLIENDQISRNHAVLTYNFVTSLSQDTEPTLTIKNSKYGT 76

QY 53 VINKLVKVKQTCPLQTCG 71
Db 77 FVNEEKMQTGLSCLTKTGD 95

RESULT 11
Q88981 PRELIMINARY; PRT; 751 AA.
AC Q88981;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Nibrin.
GN NBN.

Db 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVV 90

Qy 61 KKQTCPLQTGDVI 73

Db 91 KKQTCPLQTGDVI 103

RESULT 5

Q8BWH4 PRELIMINARY; PRT; 254 AA.

AC Q8BWH4

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Similar to cell cycle checkpoint protein CHFR (Fragment).

GN 5730484M2ORIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium.

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

DR EMBL; AK052473; BAC35008.1; -.

DR MGD; MGI:2444898; 5730484M2ORIK.

DR InterPro; IPR000253; FHA.

DR InterPro; IPR008984; SWAD_FHA.

DR Pfam; PF00498; FHA; 1.

DR SMART; SM00240; FHA; 1.

DR PROSITE; PS00006; FHA_DOMAIN; 1.

FT NON TER 254 254

SQ SEQUENCE 254 AA; 27819 MW; 0340571A2BAEB6EF CRC64;

Query Match 86.2%; Score 326; DB 11; Length 254;

Best Local Similarity 83.6%; Pred. No. 1.6e-30;

Matches 61; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVV 60

Db 31 ILLWKREWTIGRRGCDLSPFNKLVSGDHCKLTVDISGEVLTEDTSTNGTVINKLVV 90

Qy 61 KKQTCPLQTGDVI 73

Db 91 KKQTYPLQSGDII 103

RESULT 6

Q810L3 PRELIMINARY; PRT; 664 AA.

AC Q810L3

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE RIKEN cDNA 5730484M20 gene.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Limb;

RA Strausberg R.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC049792; AAH49792.1; -.

DR InterPro; IPR000253; FHA.

DR InterPro; IPR008984; SWAD_FHA.

DR InterPro; IPR001841; Znf_Ring.

DR Pfam; PF00498; FHA; 1.

DR Pfam; PF00097; Zf-C3HC4; 1.

DR SMART; SM00240; FHA; 1.

DR SMART; SM00184; RING; 2.

DR PROSITE; PS00006; FHA_DOMAIN; 1.

DR PROSITE; PS00518; ZF_RING_1; 1.

DR PROSITE; PS00089; ZF_RING_2; 1.

SQ SEQUENCE 664 AA; 73871 MW; D651BE3E463DEBB6 CRC64;

Query Match 86.2%; Score 326; DB 11; Length 664;

Best Local Similarity 83.6%; Pred. No. 4.9e-30;

Matches 61; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVV 60

Db 31 ILLWKREWTIGRRGCDLSPFNKLVSGDHCKLTVDISGEVLTEDTSTNGTVINKLVV 90

Qy 61 KKQTCPLQTGDVI 73

Db 91 KKQTYPLQSGDII 103

RESULT 7

Q8BJZ9 PRELIMINARY; PRT; 663 AA.

AC Q8BJZ9

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Similar to cell cycle checkpoint protein CHFR.

GN 5730484M2ORIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Body;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium.

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

DR EMBL; AK077629; BAC36912.1; -.

DR MGD; MGI:2444898; 5730484M2ORIK.

DR InterPro; IPR000253; FHA.

DR InterPro; IPR008984; SWAD_FHA.

DR InterPro; IPR001841; Znf_Ring.

DR Pfam; PF00498; FHA; 1.

DR Pfam; PF00097; Zf-C3HC4; 1.

DR SMART; SM00240; FHA; 1.

DR SMART; SM00184; RING; 2.

DR PROSITE; PS00006; FHA_DOMAIN; 1.

DR PROSITE; PS00518; ZF_RING_1; 1.

DR PROSITE; PS00089; ZF_RING_2; 1.

SQ SEQUENCE 663 AA; 73855 MW; 42BD653DCB4ADF45 CRC64;

Query Match 84.9%; Score 321; DB 11; Length 663;

Best Local Similarity 82.2%; Pred. No. 1.9e-29;

Matches 60; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVV 60

Db 31 ILLWKREWTIGRRGCDLSPFNKLVSGDHCKLTVDISGEVLTEDTSTNGTVINKLVV 90

Qy 61 KKQTCPLQTGDVI 73

Db 91 KKQTYPLHSGDII 103

RESULT 8

Q87PU0

QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 60
DB 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 90
QY 61 KKQTCPLQTGDVI 73
DB 91 KKQTCPLQTGDVI 103

RESULT 2
ID Q96SL3 PRELIMINARY; PRT; 652 AA.
AC Q96SL3 (T-REMBLrel. 19, Created)
DT 01-DEC-2001 (T-REMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (T-REMBLrel. 25, Last annotation update)
DE 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ14781.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RP "REDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK027687; BR555237.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 652 AA; 72030 MW; 572F2CE6D1743D80 CRC64;

Query Match 100.0%; Score 378; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.1e-36;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 60
DB 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 90
QY 61 KKQTCPLQTGDVI 73
DB 91 KKQTCPLQTGDVI 103

RESULT 3
ID Q96EP1 PRELIMINARY; PRT; 652 AA.
AC Q96EP1 (T-REMBLrel. 19, Created)
DT 01-DEC-2001 (T-REMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (T-REMBLrel. 25, Last annotation update)
DE 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Straussberg R.;

Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; BC012072; AAH12072.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 652 AA; 72058 MW; 1A50773CEA679F07 CRC64;

Query Match 100.0%; Score 378; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 3.1e-36;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 60
DB 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 90
QY 61 KKQTCPLQTGDVI 73
DB 91 KKQTCPLQTGDVI 103

RESULT 4
ID Q9NRT4 PRELIMINARY; PRT; 664 AA.
AC Q9NRT4 (T-REMBLrel. 15, Created)
DT 01-OCT-2000 (T-REMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (T-REMBLrel. 25, Last annotation update)
DE 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)
DE Cell cycle checkpoint protein CHFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2038685; PubMed=10935642;
RA Scolnick D.M., Halazonetis T.D.;
RT "Chfr defines a mitotic stress checkpoint that delays entry into
RT metaphase."
RL Nature 406:430-435(2000).
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AF170724; AAF91084.1; -.
DR PDB; ILGP; 07-AUG-02.
DR PDB; ILGQ; 07-AUG-02.
DR Genew; HGNC:20455; CHFR.
DR GO; GO:0007093; P:mitotic checkpoint; TAS.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 664 AA; 73386 MW; 141A1E7FEFAE36A2 CRC64;

Query Match 100.0%; Score 378; DB 4; Length 664;
Best Local Similarity 100.0%; Pred. No. 3.2e-36;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 60
DB 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 90
QY 61 KKQTCPLQTGDVI 73
DB 91 KKQTCPLQTGDVI 103

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: May 7, 2004, 14:40:47 ; Search time 10.5607 Seconds
(without alignments)
2180.991 Million cell updates/sec

Title: US-10-048-046-2_COPY_31_103
Perfect score: 378
Sequence: 1 VLLKREWTIGRRGCDLSF.....INKLKVVKQTCPLQGDVI 73

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	378	100.0	623	4 Q9NVDS	Q9NVDS homo sapien
2	378	100.0	652	4 Q96SL3	Q96SL3 homo sapien
3	378	100.0	652	4 Q96EP1	Q96EP1 homo sapien
4	378	100.0	664	4 Q9NR74	Q9NR74 homo sapien
5	326	86.2	254	11 Q8BWH4	Q8BWH4 mus musculus
6	326	86.2	664	11 Q810L3	Q810L3 mus musculus
7	321	84.9	663	11 Q8BZ29	Q8BZ29 mus musculus
8	107.5	28.4	504	16 Q87PU0	Q87PU0 vibrio para
9	88	23.3	159	16 Q97LQ0	Q97LQ0 clostridium
10	87	23.0	546	11 Q8CCY0	Q8CCY0 mus musculus
11	87	23.0	751	11 Q88981	Q88981 mus musculus
12	87	23.0	751	11 Q9R1X1	Q9R1X1 mus musculus
13	87	23.0	751	11 Q811I6	Q811I6 mus musculus
14	87	23.0	751	11 Q9R207	Q9R207 mus musculus
15	86	22.8	176	16 Q8G5Q4	Q8G5Q4 bifidobacte
16	83	22.0	243	16 Q7UWU2	Q7UWU2 rhodospirell

17	82.5	21.8	729	10 Q9FMB8	Q9FMB8 arabidopsis
18	79.5	21.0	497	16 Q91751	Q91751 pseudomonas
19	79.5	21.0	1236	5 Q9NED7	Q9NED7 leishmania
20	79	20.9	250	16 Q7UGY0	Q7UGY0 rhodospirell
21	79	20.9	345	4 Q9UBH7	Q9UBH7 homo sapien
22	79	20.9	345	4 Q9Y242	Q9Y242 homo sapien
23	79	20.9	345	4 Q9EUM2	Q9EUM2 homo sapien
24	79	20.9	345	6 Q7YR48	Q7YR48 pan troglod
25	79	20.9	359	4 Q13176	Q13176 homo sapien
26	79	20.9	359	16 Q98IK6	Q98IK6 rhizobium 1
27	77.5	20.5	460	16 Q98P85	Q98P85 bradyrhizob
28	77.5	20.5	528	3 Q874N8	Q874N8 candida gla
29	76.5	20.2	517	13 Q918V3	Q918V3 xenopus lae
30	76	20.1	2541	5 Q19663	Q19663 caenorhabdi
31	74.5	19.7	263	11 Q99KJ5	Q99KJ5 mus musculu
32	74.5	19.7	263	11 Q8BPA3	Q8BPA3 mus musculu
33	74.5	19.7	517	13 Q98TW0	Q98TW0 xenopus lae
34	74.5	19.7	689	10 Q9FN27	Q9FN27 oryza sativ
35	73.5	19.4	526	3 Q874M7	Q874M7 klyveromyc
36	73	19.3	144	16 Q990Y0	Q990Y0 clostridium
37	73	19.3	503	13 Q90ZY5	Q90ZY5 brachydanio
38	71.5	18.9	398	16 Q87HC2	Q87HC2 vibrio para
39	71	18.8	503	13 Q803B2	Q803B2 brachydanio
40	70.5	18.7	168	5 Q86DY9	Q86DY9 schistosoma
41	70.5	18.7	230	16 Q9KBZ7	Q9KBZ7 bacillus ha
42	70.5	18.7	301	3 Q74388	Q74388 schizosacch
43	70	18.5	254	2 Q9S475	Q9S475 myxococcus
44	70	18.5	388	2 Q52176	Q52176 myxococcus
45	68.5	18.1	157	16 P74513	P74513 synectocyst

ALIGNMENTS

RESULT 1
Q9NVDS PRELIMINARY; PRT; 623 AA.
AC Q9NVDS; Q9NVDS (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ10796.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.,
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK001658; BAA91817.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Fing.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
DR Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 623 AA; 69204 MW; 45186D33DAE52711 CRC64;

Query Match 100.0%; Score 378; DB 4; Length 623;
Best Local Similarity 100.0%; Pred. No. 3e-36;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 T-----SGTVI-----NKLKVKVKQTC 65
 DB 129 TNHFEIAGTTVKSQWFKSTTYITLDTQC 156

RESULT 14
 TMU CORGL
 ID TMU CORGL STANDARD; PRT; 365 AA.
 AC Q8NR24;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable tRNA (5-methylaminomethyl-2-thiouridylylate)-methyltransferase
 DE (EC 2.1.1.61).
 GN TMU OR CGL1240.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-homocysteine + tRNA containing 5-methylaminomethyl-2-thiouridylylate.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the tmU family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AP005277; BAB98633.1; -;
 DR HAMAP; MF 00144; -; 1.
 DR InterPro; IPR004506; TmU.
 DR Pfam; PF03054; tRNA.Me.trans; 1.
 DR TIGRFAMs; TIGR00420; tmU; 1.
 DR Transferase; Methyltransferase; tRNA processing; Complete proteome.
 SQ SEQUENCE 365 AA; 39237 MW; 21F8EE1B2B3BDC63 CRC64;

Query Match 16.0%; Score 60.5; DB 1; Length 365;
 Best Local Similarity 29.6%; Pred. No. 18;
 Matches 21; Conservative 14; Mismatches 23; Indels 13; Gaps 4;

QY 7 EWTGRRRGCDLSPSKNLVSGDHCRIV--VDEKSGQVTL---EDTSTSGTVINKLVKVK 61
 DB 237 EFTIGQRKGLDIKAPA-----ADGRPRYVTDIDAKGTVTGTRENLKISTTHADRLKPLH 292

QY 62 KQTCPLQGTGV 72
 DB 293 ----PAMDQOI 299

RESULT 15
 Y634 CHLMU
 ID Y634 CHLMU STANDARD; PRT; 353 AA.
 AC Q9PK39;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein TC0634 precursor.
 GN Chlamydia muridarum.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=833560;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=MoPn / Ni9G;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 CC -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN1058/CT355/TC0634
 CC FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE002332; AAF39463.1; -;
 DR PIR; E81680; E81680.
 DR TIGR; TC0634; -;
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 24
 FT CHAIN 25 353
 FT CHAIN 25 353
 SQ SEQUENCE 353 AA; 39840 MW; 3048C59F5B7BB90E CRC64;

Query Match 15.9%; Score 60; DB 1; Length 353;
 Best Local Similarity 32.2%; Pred. No. 20;
 Matches 19; Conservative 10; Mismatches 16; Indels 14; Gaps 3;

QY 15 GCDLSPSKNLVSGD-----HCRIV--VDEKSGQVTLTSTSGTVINKLVKVKQTC 66
 DB 24 GYEGSPSSLEQNPQSGVAIHNRVLFKIDETVTTLD-----VIHKLNIIFYSTCP 75

Search completed: May 7, 2004, 14:46:32
 Job time : 4.46674 secs

```
Matches 22; Conservative 9; Mismatches 34; Indels 12; Gaps 3;
QY 8 WTGRRGCDLSPSKNL-----VSGDHCRIVD---EKSQVTLDTSTSGVINK 56
Db 118 W-FGRDKSCYCFDGLLRTRDKYRTYKSKHFRIFREMGPKNFYIYVIEDHSGNGTFVNT 176
QY 57 LKVVKTKCPLOQGDVI 73
Db 177 ELGKGRKCPLSNNSEI 193

RESULT 12
VGLM_BUNYW
ID_VGLM_BUNYW STANDARD; PRT; 1433 AA.
AC P04505;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE M polyprotein precursor [Contains: Glycoprotein G2; Nonstructural
DE protein NS-M; Glycoprotein G1].
GN M.
OS Bunyamwera virus.
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.
OX NCBI_TaxID=35304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86098655; PubMed=3753629;
RA Lees J.F., Pringle C.R., Elliott R.M.;
RT "Nucleotide sequence of the Bunyamwera virus M RNA segment:
RT conservation of structural features in the Bunyamwera glycoprotein
RT gene product.";
RL Virology 148:11-14 (1986).
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins
CC including nonstructural protein NS-M, glycoprotein G1, and
CC glycoprotein G2.
CC -1- SIMILARITY: Belongs to the bunyaviruses M polyprotein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M11852; AAA42777.1; -.
CC PIR; A04101; GNVUBW.
CC InterPro; IPR005167; Bunya_G1.
CC InterPro; IPR005168; Bunya_G2.
CC Pfam; PF03557; Bunya_G1; 1.
CC Pfam; PF03563; Bunya_G2; 1.
CC Polyprotein; Glycoprotein; Transmembrane; Nonstructural protein;
CC Signal.
CC SIGNAL 1 16
CC CHAIN 17 1433 M POLYPROTEIN.
CC CHAIN 17 302 GLYCOPROTEIN G2.
CC CHAIN 303 477 NONSTRUCTURAL PROTEIN NS-M.
CC CHAIN 478 1433 GLYCOPROTEIN G1.
CC CHAIN 1382 1409 POTENTIAL.
CC TRANSMEM 1382 1409 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 1169 1169 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 1433 AA; 162077 MW; CD61ABDE782018E0 CRC64;

Query Match 16.1%; Score 61; DB 1; Length 1433;
Best Local Similarity 23.9%; Pred. No. 70;
Matches 21; Conservative 9; Mismatches 30; Indels 28; Gaps 3;
QY 6 REWTIGRRGC-----DLSPFNKLVSGDHCRIVVDEKSGV-----T 43
Db 72 RHWITDKACNPVVTAGGSINVIKLVNLTNRNVCTGDCITVDRKNAQIIFQDK 131
```

```
QY 44 LEDTSTSGTVI-----NKLKVVKKQTC 65
Db 132 LNHFVETGTTTSTGWFKSKASVTLDRIC 159

RESULT 13
VGLM_BUNL7
ID_VGLM_BUNL7 STANDARD; PRT; 1441 AA.
AC P09612;
DT 01-NAR-1989 (Rel. 10, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE M polyprotein precursor [Contains: Glycoprotein G2; Nonstructural
DE protein NS-M; Glycoprotein G1].
GN M.
OS Bunyavirus La Crosse (isolate L74).
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.
OX NCBI_TaxID=11578;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88089508; PubMed=3694177;
RA Grady L.J., Sanders M.L., Campbell W.P.;
RT "The sequence of the M RNA of an isolate of La Crosse virus.";
RL J. Gen. Virol. 68:3057-3071 (1987).
RN [2]
RP SEQUENCE OF 1-46 FROM N.A.
RX MEDLINE=82216937; PubMed=7086954;
RA Clerx-Van Haaster C.M., Akashi H., Auperin D.D., Bishop D.H.L.;
RT "Nucleotide sequence analyses and predicted coding of bunyavirus
RT genome RNA species.";
RL J. Virol. 41:119-128 (1982).
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins
CC including nonstructural protein NS-M, glycoprotein G1, and
CC glycoprotein G2.
CC -1- SIMILARITY: Belongs to the bunyaviruses M polyprotein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10370; BAA01201.1; -.
CC PIR; J02231; -. NOT_ANNOTATED_CDS.
CC PIR; A29377; GNVULC.
CC InterPro; IPR005167; Bunya_G1.
CC InterPro; IPR005168; Bunya_G2.
CC Pfam; PF03557; Bunya_G1; 1.
CC Pfam; PF03563; Bunya_G2; 1.
CC Polyprotein; Glycoprotein; Transmembrane; Nonstructural protein;
CC Signal.
CC SIGNAL 1 13
CC CHAIN 14 1441 M POLYPROTEIN.
CC CHAIN 22 299 GLYCOPROTEIN G2.
CC CHAIN 300 473 NONSTRUCTURAL PROTEIN NS-M.
CC CHAIN 474 1441 GLYCOPROTEIN G1.
CC CHAIN 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 1177 1177 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CONFLICT 40 40 C -> S (IN REF. 2).
CC SEQUENCE 1441 AA; 162540 MW; 50973CE30973C55B CRC64;

Query Match 16.1%; Score 61; DB 1; Length 1441;
Best Local Similarity 22.7%; Pred. No. 70;
Matches 20; Conservative 12; Mismatches 28; Indels 28; Gaps 3;
QY 6 REWTIGRRGC-----DLSPFNKLVSGDHCRIVVDEKSGV-----T 48
Db 69 RKWLVDWHDRCFKPKKIVGSHINVIKLVNLTNRNVCTGDCITVDRKNAQIIFQDK 128
```

```
-- -- -1- SIMILARITY: Contains 1 fork-head domain.
```

```
-- -- -1- SIMILARITY: Contains 1 FHA domain.
```

```
-
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to licensee@isb-sib.ch).
```

```
-- --
```

```
EMBL; L38848; AAA6093B.1; --  
EMBL; Z38059; CAA86147.1; --  
PIR; S48403; S48403.  
HSP; Q63245; 2HFH.  
GeneOnline; I39666; -.  
TRANSFAC; T03303; -.  
SGD; S0001393; FKHI.  
GO; GO:0005634; C:nucleus; IDA.  
DR GO; GO:0006347; P:chromatin silencing at HML and HMR (sensu S. ..; IGI.  
DR GO; GO:0007124; P:pseudohyphal growth; IGI.  
DR GO; GO:0000074; P:regulation of cell cycle; IGI.  
InterPro; IPRO00253; FHA.  
DR InterPro; IPRO01766; TF_Fork_head.  
DR Pfam; PF00498; FHA; 1.  
DR Pfam; PF00250; Fork head; 1.  
PRINTS; PR00053; FORKHEAD.  
DR PRODOM; PD000425; TF_Fork_head; 1.  
SMART; SMART; SM00339; FH; 1.  
SMART; SMART; SM00240; FHA; 1.  
DR PROSITE; PSS0006; FHA DOMAIN; 1.  
DR PROSITE; PS00657; FORK HEAD_1; 1.  
DR PROSITE; PS00658; FORK HEAD_2; 1.  
DR PROSITE; PSS00039; FORK_HEAD_3; 1.  
KW DNA-binding; Nuclear protein.  
FT DNAMEIN 76  
FT DOMAIN 142  
FT BINDING 302  
FT SEQUENCE 484 AA; 53490 MW; 7159073EB979C489 CRC64;  
-- --
```

```
Query Match          16.1%; Score 61; DB 1; Length 484;  
Best Local Similarity 24.7%; Pred.No. 21;  
Matches      22; Conservative     15; Mismatches    34; Indels   18; Gaps       2;
```

```
QY        3 LRKEWTIGR-----RRGCLDPFPNKLVSGDHCRIIVVDEKSQGVTLED 47  
         :::|||::|::|::|::|::|::|::|::|::|::|:  
Db        71 VQLKVTVIGRTSLNLNAVPGTWVKXIDDLGPAKIVSRKHAAIRFNLSGSWELQIF 130  
         ::|::|::|::|::|::|::|::|::|::|::|:
```

```
QY        48 STGGTVINKLVKKQTCP---LTGTGV 73  
         :::|::|::|::|::|::|::|::|::|::|:  
Db       131 GRNGAKYNFRIPRGPSPTVLQSGLII 159  
         :::|::|::|::|::|::|::|::|::|::|:
```

```
RESULT 11  
CHK2_MOUSE  
ID   CHK2_MOUSE STANDARD; PRT; 546 AA.  
AC Q9Z265;  
DT 30-MAY-2000 (Rel. 39, Created)  
DD 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 15-MAR-2004 (Rel. 43, Last annotation update)  
DN Serine/threonine-protein kinase Chk2 (EC 2.7.1.37).  
GN CHK2 OR CHK2 OR RAD53.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
CX NCBI_Taxid=10090;  
RN [1]  
RP RNCB_SEQUENCE FROM N.A.  
RX MEDLINE=98055399; PubMed=9836640;  
RA Matsuo S., Huang M., Ellledge S.J.;  
RT "Linkage of ATM to cell cycle regulation by the Chk2 protein kinase.";  
RL Science 282:1893-1897(1998).  
RN [2]
```

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195(2000).
RN [4]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminck J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review,"
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RL [5]
RN SEQUENCE FROM N.A. (ISOFORM SHORT).
RP STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarini H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RA "A *Drosophila* full-length cDNA resource,"
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -!- FUNCTION: May have a role in germ line establishment.
CC -!- SUBCELLULAR LOCATION: Nuclear; speckled subnuclear compartment.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=061267-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=061267-2; Sequence=VSP 004865;
CC -!- TISSUE SPECIFICITY: In stage 3 embryos, both isoforms are
CC expressed in both somatic and pole cell nuclei. Expression in pole
CC cell nuclei is sustained until stage 9 and weakly expressed after
CC pole cell invagination into the abdominal cavity.
CC -!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically in
CC adult females. Levels of the long isoform remain fairly constant
CC from ovaries to embryos, the levels of short isoform decrease
CC dramatically.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CDS1
CC subfamily.
CC -!- SIMILARITY: Contains 1 FHA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB007821; BAA28755.1; -
CC EMBL: AB007822; BAA28756.1; -
CC EMBL: U87984; AAB49642.1; -
CC EMBL: AE003665; AAF3867.2; -
CC EMBL: AE003665; AAN1062.1; -
CC EMBL: AY070549; AAL48020.1; -
CC HSSP: Q63450.1A06.
CC FlyBase: FBgn019686; lok.
CC GO: GO:0005634; C:nucleus; IEP.

GO: GO:0004674; P:protein serine/threonine kinase activity; NAS.
GO: GO:0008630; P:DNA damage response, signal transduction re.; IMP.
GO: GO:0000077; P:DNA damage response, signal transduction re.; IMP.
GO: GO:0007281; P:germ-cell development; IEP.
GO: GO:0006468; P:protein amino acid phosphorylation; NAS.
InterPro: IPR000253; FHA.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_Thr_kin_A.
DR InterPro: IPR002290; Ser_Thr_kinase.
DR InterPro: IPR008984; SMAD_FHA.
DR Pfam: PF00498; FHA; 1.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM002240; FHA; 1.
DR SMART: SM00220; S_TK; 1.
DR PROSITE: PS50006; FHA_DOMAIN; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
KW Transferase: Serine/threonine-protein kinase; ATP-binding;
KW Nuclear protein; Alternative splicing.
FT DOMAIN 69 129 FHA
FT DOMAIN 174 441 PROTEIN KINASE
FT NP_BIND 180 188 ATP (BY SIMILARITY).
FT BINDING 203 203 ATP (BY SIMILARITY).
FT ACT_SITE 303 303 BY SIMILARITY.
FT VARSPLIC 46 62 Missing (in isoform Short).
FT /FTID=VSP 004865.
SQ SEQUENCE 476 AA; 54261 MW; 58D583E015C4B626 CRC64;
Query Match. 16.3%; Score 61.5; DB 1; Length 476;
Best Local Similarity 25.9%; Pred. No. 18;
Matches 22; Conservative 11; Mismatches 33; Indels 19; Gaps 2;
QY 3 LRKREWTIGRRGCDLSPSNKL-----VSGDHCRIVWDEKSGQVLEDTLS 48
DB 64 LNDSEFTAGRGANDLILTLNLPKILTRISKVFIKRCAN-----ELTNPVVIQDLS 118
QY 49 TSGTVINKLVVYKQTCPLQGTGVI 73
DB 119 RGTGFVYNEKIGTNRMRILKNDVI 143
RESULT 10
FKH1 YEAST STANDARD; PRT; 484 AA.
ID FKH1 YEAST
AC P40456;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fork head protein homolog 1.
GN FKH1 OR YIL131C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C;
RA Zhu G., Davis T.N.;
RL Submitted (XXX-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moule S., Odell C., Pearson D., Rajadream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
RA "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX,"
RT Nature 387:84-87(1997).
RL CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).

Lyne M.H., Bryant J.A., Aves S.J.;
 Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 [2]
 SEQUENCE FROM N.A.
 STRAIN-972;
 MEDLINE-21648401; PubMed-11859360;
 Wood V.G., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 Rutherford K., Rutter S., Saunders R., Squares S., Stevens K.,
 Skelton J., Simmonds M., Squares R., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 Wray K., Taylor R.G., Volckaert G., Aert R., Robben J., Grynolprez B.,
 Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 Cabel C., Fuchs M., Pritz C., Holzer E., Moestl D., Hilbert H.,
 Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 Lucas M., Rochet M., Gaillard C., Tallada V.A., Garzon A., Thode G.,
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 "The genome sequence of *Schizosaccharomyces pombe*.";
 Nature 415:871-880(2002).
 -!- FUNCTION: Probable protein kinase required for meiotic
 recombination.
 -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 -!- SIMILARITY: Contains 1 FHA domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 or send an email to license@sib-sib.ch).

 EMBL; Z71478; CAA96101.1; --
 EMBL; Z85836; CAB11196.1; --
 PIR; T43420; T43420.
 HSSP; P00518; 1PKH.
 GenDB_Spombe; SPAC14C4.03; --
 InterPro; IPR000253; FHA.
 InterPro; IPR000719; prot_kinase.
 InterPro; IPR008271; Ser_thr_pkin_AS.
 InterPro; IPR002290; Ser_thr_pkinase.
 InterPro; IPR008984; SMAD_FHA.
 InterPro; IPR001245; Tyr_pkinase.
 Pfam; PF00498; FHA; 1.
 Pfam; PF00069; pkinase; 1.
 PRINTS; P00109; TYRKINASE.
 ProDom; PD000001; prot_kinase; 1.
 SMART; SM00240; FHA; 1.
 SMART; SM00220; S_TKC; 1.
 PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 PROSITE; PS00066; FHA_DOMAIN; 1.
 Transferase; Serine/threonine-protein kinase; ATP-binding; Meiosis.
 DOMAIN 62 116 FHA
 DOMAIN 160 421 PROTEIN_KINASE
 NP_BIND 166 174 ATP (BY SIMILARITY).
 ACT_SITE 281 281 BY SIMILARITY.
 BINDING 189 189 ATP (BY SIMILARITY).
 FT
 KW
 FT

SQ SEQUENCE 445 AA; 51151 MW; 6D62D79E9A1B45B0 CRC64;
 Query Match 16.8%; Score 63.5; DB 1; Length 445;
 Best Local Similarity 27.8%; Pred.No.10;
 Matches 20; Conservative 14; Mismatches 27; Indels 11; Gaps 3;

 QY 9 TIGRRRGCD--LSFSPNKLVSGDHCR-----VWDEKSGQVTLEDTSTSGTVINKLXVK 61
 ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
 Db 63 SVGRSNTCNVLQLQFTA----SYKHFRVYSVLIDDDMDPLVYCDDSSNGGFLNHRLLGX 118
 :||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
 QY 62 KQTCPLOTGVPI 73
 :||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
 Db 119 GNSVLLSDGDIL 130
 :||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

 RESULT 9
 LOK DROME STANDARD; PRT; 476 AA.
 ID LOK DROME STANDARD; PRT; 476 AA.
 AC O61267; O61268; P91876; Q8S233;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ovarian-specific serine/threonine-protein kinase Lok (EC 2.7.1.-)
 (loki protein) (GNM).
 OS LOK OR CGI0895.
 GN Drosohila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT), FUNCTION, SUBCELLULAR
 RP LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC STRAIN=Canton-S; TISSUE=Embryo;
 RX MEDLINE=98175876; PubMed=9507063;
 RA Oishi I., Sugiyama S., Otani H., Yamamura H., Nishida Y., Minami Y.;
 RT "A novel drosophila nuclear protein serine/threonine kinase expressed
 RL in the germline during its establishment."; Mech. Dev. 71:49-63(1998).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM SHORT).
 RP STRAIN=Oregon-R; TISSUE=Ovary;
 RC Larochelle S., Suter B.;
 RX Identification of a novel ovarian specific protein kinase.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DDDB databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Berkelley;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams W.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yeandle M.D., Zhang Q., Chen J.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Benson K.Y., Berno P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Curtis J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paulis B., Delcher A., Deng Z., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Fleischmann W.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hostin N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M.A., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mosherfi.,

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1] SEQUENCE FROM N.A. (ISOFORM 2).
RN TISSUE=Brain;
RC MEDLINE=97177144; PubMed=9024696;
RX Kim E., Naisbitt S., Haueh Y.-P., Rao A., Rothschild A., Craig A.M.,
RA Sheng M.;
RT "GKAP, a novel synaptic protein that interacts with the guanylate
RT kinase-like domain of the PSD-95/SAP90 family of channel clustering
RT molecules.";
RL J. Cell Biol. 136:669-678 (1997).
[2] SEQUENCE FROM N.A. (ISOFORM 3).
RN MEDLINE=98089008; PubMed=9428732;
RX Kawashima N., Takamiya K., Sun J., Kitabatake A., Sobue K.;
RA "Differential expression of isoforms of PSD-95 binding protein
RT (GKAP/SAPAP1) during rat brain development.";
RL FEBS Lett. 418:301-304 (1997).
[3] SEQUENCE FROM N.A. (ISOFORM 1).
RN TISSUE=Brain;
RC MEDLINE=97277335; PubMed=9115257;
RX Takeuchi M., Hata Y., Hirao K., Toyoda A., Irie M., Takai Y.;
RA "SAPAPs. A family of PSD-95/SAP90-associated proteins localized at
RT postsynaptic density.";
RL J. Biol. Chem. 272:11943-11951 (1997).
[4] INTERACTION WITH DLG4 AND SHANK PROTEINS.
RN MEDLINE=99458653; PubMed=10527873;
RX Boeckers T.M., Winter C., Smalla K.-H., Kreutz M.R., Bockmann J.,
RA Seidenbecher C., Garner C.C., Gundelfinger E.D.;
RT "proline-rich synapse-associated proteins ProsAP1 and ProsAP2 interact
RT with synaptic proteins of the SAPAP/GKAP family.";
RL Biochem. Biophys. Res. Commun. 264:247-252 (1999).
[5] INTERACTION WITH DLG4 AND SHANK1.
RN TISSUE=Brain;
RC MEDLINE=99419021; PubMed=10488079;
RX Yao I., Hata Y., Hirao K., Deguchi M., Ide N., Takeuchi M., Takai Y.;
RA "Synapin, a novel neuronal protein interacting with synapse-associated
RT protein 90/postsynaptic density-95-associated protein.";
RL J. Biol. Chem. 274:27463-27466 (1999).
[6] INTERACTION WITH SHANK PROTEINS, AND MUTAGENESIS ALA-990 AND ALA-992.
RN MEDLINE=99360650; PubMed=10433268;
RX Naisbitt S., Kim E., Tu J.-C., Xiao B., Sala C., Valtchanoff J.,
RA Weinberg R.J., Worley P.F., Sheng M.;
RT "Shank, a novel family of postsynaptic density proteins that binds to
RT the NMDA receptor/PSD-95/GKAP complex and cortactin.";
RL Neuron 23:569-582 (1999).
[7] INTERACTION WITH AIP1.
RN MEDLINE=98361985; PubMed=9694864;
RX Hirao K., Hata Y., Ide N., Takeuchi M., Irie M., Yao I., Deguchi M.,
RA Toyoda A., Sudhof T.C., Takai Y.;
RT "A novel multiple PDZ domain-containing molecule interacting with
RT N-methyl-D-aspartate receptors and neuronal cell adhesion proteins.";
RL J. Biol. Chem. 273:21105-21110 (1998).
-!- FUNCTION: May play a role in the molecular organization of
CC synapses and neuronal cell signaling. Could be an adapter protein
CC linking ion channel to the subsynaptic cytoskeleton. May induce
CC enrichment of PSD-95/SAP90 at the plasma membrane.
CC -!- SUBUNIT: Interacts with DLG1, DLG4/PSD-95, APC and AIP1. The C-
CC terminus binds the PDZ domain of the SHANK1, SHANK2 and SHANK3. Is
CC part of a complex with DLG4/PSD-95 and SHANK1, SHANK2 or SHANK3.
CC -!- SUBCELLULAR LOCATION: Membrane-associated.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=p97836-1; Sequence=Displayed;
CC Name=2;
CC IsoId=p97836-2; Sequence=VSP_006007, VSP_006008, VSP_006009,
CC

```

DR HSP; Q63450; 1A06.
DR Germline; 140343; -.
DR SGD; S0002259; DUN1.
DR GO; GO:0004572; F:protein kinase activity; IDA.
DR GO; GO:0000077; P:DNA damage response, signal transduction re. .; IMP.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR002230; Ser_thr_kinase.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE DM; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Nuclear protein; Serine/threonine-protein kinase; DNA damage.
FT DOMAIN 56 112
FT PROTEIN KINASE.
FT NP_BIND 200 480
FT ATP (BY SIMILARITY).
FT BINDING 206 214
FT ATP (BY SIMILARITY).
FT BINDING 229 229
FT ATP (BY SIMILARITY).
FT ACT_SITE 328 328
FT K-R: LOSS OF KINASE FUNCTION.
FT MUTAGEN 328 328
FT D-A: LOSS OF KINASE FUNCTION.
FT MUTAGEN 328 328
FT SEQUENCE 513 AA; 58632 MW; 4990F24F024702D7 CRC64;

Query Match 19.0%; Score 72; DB 1; Length 513;
Best Local Similarity 35.7%; Pred. No. 1.3;
Matches 25; Conservative 10; Mismatches 29; Indels 6; Gaps 3;

QY 9 TIGRRCDD-LAPPNKLVSVDHCRIVDEKSGQ---VLEDTSTGTINKLVKVKQ 63
DB 57 TIGRSCDVLSEPDISTFHAFLQMDVDFQRLNIVDKSRNGFINGRLVKD 116
QY 64 TGPLQTDVI 73
DB 117 YI-LKNGDRI 125

RESULT 4
ID PGK MYCPN STANDARD; PRT; 409 AA.
AC P78019;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK OR MPN429 OR MP412.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae"; Res. 24:4420-4449(1996).
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -1- PATHWAY: Second phase of glycolysis; second step.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the phosphoglycerate kinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000040; AAB96060.1; -.
DR PIR; S73738; S73738.
DR HSP; P36204; 1YPE.
DR HAMAP; MF_00145; 1.
DR InterPro; IPR001576; PGK.
DR Pfam; PF00162; PGK; 1.
DR PRINTS; PR00477; PEGLYCKINASE.
DR PROSITE; PS00111; POLYGLYCERATE KINASE; 1.
KW Transferase; Kinase; Glycolysis; Complete proteome.
SQ SEQUENCE 409 AA; 44212 MW; 6AF230188D398731 CRC64;

Query Match 18.8%; Score 71; DB 1; Length 409;
Best Local Similarity 29.2%; Pred. No. 1.3;
Matches 19; Conservative 14; Mismatches 20; Indels 12; Gaps 2;

QY 19 SFPSNKLVSVDHCRIV-----VDEKSGQVLTSTGTINKLVKVKQTCLQ 68
DB 41 SLDITKILLGHCKIVLLSHLSRVKSLDDKKGKSLQPVASA--LQNLKNTKVHFCPEN 98
QY 69 TGDVI 73
DB 99 TGDVI 103

RESULT 5
ID DLPI_HUMAN STANDARD; PRT; 977 AA.
AC O14490; O14489; P78335;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Disks large-associated protein 1 (DAP-1) (guanylate kinase-associated
DE protein) (HGKAP) (SAP90/PSD-95-associated protein 1) (SAPAP1) (PSD-
DE 95/SAP90 binding protein 1).
GN DLGAP1 OR GKAP OR DAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=97431353; PubMed=9286858;
RA Satoh K., Yanai H., Senda T., Kohu K., Nakamura T., Okumura N.,
RA Matsumine A., Kobayashi S., Toyoshima K., Akiyama T.;
RT "DAP-1, a novel protein that interacts with the guanylate kinase-like
RT domains of hDLG and PSD-95.";
RL Genes Cells 2:415-424(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE=9717144; PubMed=9024696;
RA Kim E., Naisbitt S., Hsueh Y.-P., Rao A., Rothschild A., Craig A.M.,
RA Sheng M.;
RT "GKAP, a novel synaptic protein that interacts with the guanylate
RT kinase-like domain of the PSD-95/SAP90 family of channel clustering
RT molecules.";
RL J. Cell Biol. 136:669-678(1997).
CC -1- FUNCTION: May play a role in the molecular organization of
CC synapses and neuronal cell signaling. Could be an adapter protein
CC linking ion channel to the subsynaptic cytoskeleton. May induce
CC enrichment of PSD-95/SAP90 at the plasma membrane.
CC -1- SUBUNIT: Interacts with DLG1, DLG4/PSD-95, APC and AIP1 (By
CC similarity). Isoform 1 and isoform 2 C-terminus bind the PDZ
CC domain of SHANK1, SHANK2 and SHANK3 (By similarity). Is part of a
CC complex with DLG4/PSD-95 and SHANK1, SHANK2 or SHANK3 (By
CC similarity).
CC

```

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe I., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipakowski G.V., Usery D., Barrell B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: Has a role in the DNA replication-monitoring S/G2
CC checkpoint system. It is responsible for blocking mitosis in the
CC S phase. It monitors DNA synthesis by interacting with DNA
CC polymerase alpha and sends a signal to block the onset of mitosis
CC while DNA synthesis is in progress.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBUNIT: Interacts with rad26.
CC -1- PTM: Autophosphorylated.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CDS1
CC subfamily.
CC -1- SIMILARITY: Contains 1 FHA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X85040; CAA59410.1; -.
CC EMBL; AJ222869; CAA11019.1; -.
CC EMBL; AL109736; CAB52158.1; -.
CC PIR; S58882; S58882.
CC HSSP; Q00534; 1B17.
CC GeneDB SPombe; SPCC18B5.11c; -.
CC InterPro; IPR000253; FHA.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR008271; Ser Thr pkin AS.
CC InterPro; IPR002290; Ser Thr pkinase.
CC InterPro; IPR008984; SMAD FHA.
CC InterPro; IPR001245; Tyr pkinase.
CC Pfam; PF00498; FHA; 1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot kinase; 1.
CC SMART; SM00240; FHA; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE-ST; 1.
CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00006; FHA DOMAIN; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
CC Cell cycle; Phosphorylation.
CC DOMAIN 60 116 FHA.
CC FT 167 433 PROTEIN KINASE.
CC FT NP_BIND 173 181 ATP (BY SIMILARITY).
CC FT BINDING 196 196 ATP (BY SIMILARITY).
CC FT ACT_SITE 294 294 BY SIMILARITY.
CC FT CONFLICT 61 61 R -> G (IN REF. 1).
CC FT CONFLICT 238 238 F -> C (IN REF. 1).

SQ SEQUENCE 460 AA; 52014 MW; 4CEB963D3376DB54 CRC64;
Query Match 20.5%; Score 77.5; DB 1; Length 460;
Best Local Similarity 30.6%; Pred. No. 0.27;
Matches 22; Conservative 34; Indels 7; Gaps 2;
QY 8 WTIGRRGCDLSPPSNKLVSGDHCEIV-----VDEKSGQVLTETSTSGTVINKLVKK 61
DB 60 WFGFGHKCEVVL-NGPRVSNFHFYIYQGHNRNDSSENVLHDHSSNGTFLNPERLAK 118
QY 62 KQTCPLQTQGDVI 73
DB 119 NSRTILSNGDEI 130
RESULT 3
DUNI_YEAST
ID DUNI_YEAST STANDARD; PRT; 513 AA.
AC P39009;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA damage response protein kinase DUNI (EC 2.7.1.-).
GN DUNI OR YDL101C OR D370.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94084787; PubMed=8261511;
RA Zhou Z., Elledge S.J.;
RT "DUNI encodes a protein kinase that controls the DNA damage response
RT in yeast.";
RL Call 75:1119-1127(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97051597; PubMed=8896274;
RA Saiz J.E., Buicrago M.J., Garcia R., Revuelta J.L., del Rey F.;
RT "The sequence of a 20.3 kb DNA fragment from the left arm of
RT Saccharomyces cerevisiae chromosome IV contains the KIN28, MSS2,
RT PHO2, POL3 and DUNI genes, and six new open reading frames.";
RN Yeast 12:1077-1084(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97082507; PubMed=8923743;
RA Boskovic J., Saiz J.E., Soler-Mira A., Garcia-Cantalejo J.M.,
RA Ballesta J.P.G., Jimenez A., Remacha M.;
RT "The sequence of a 16,691 bp segment of Saccharomyces cerevisiae
RT chromosome IV identifies the DUNI, PMT1, PMT5, SRP14 and DPR1 genes,
RT and five new open reading frames.";
RL Yeast 12:1377-1384(1996).
CC -1- FUNCTION: Transducer of the DNA damage signal.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: Autophosphorylation increases in response to DNA damage.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CDS1
CC subfamily.
CC -1- SIMILARITY: Contains 1 FHA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L25549; AAA16324.1; -.
CC EMBL; X95644; CAA64912.1; -.
CC EMBL; Z74149; CAA98668.1; -.
CC PIR; S43941; S43941.

RT damage-induced transcription in yeast.";
RN Genes Dev. 8:2401-2415(1994).
RP FUNCTION
RX MEDLINE=20018334; PubMed=10550056;
RA Sanchez Y., Bachant J., Wang H., Fu F., Liu D., Tetzlaff M.,
RA Billede S.J.;
RT "Control of the DNA damage checkpoint by chk1 and rad53 protein
RT kinases through distinct mechanisms.";
RL Science 286:1166-1171(1999).
RN [6]
RN PHOSPHORYLATION.
RX MEDLINE=20031667; PubMed=10562568;
RA Pelliccioli A., Lucca C., Liberi G., Marini F., Lopes M., Plevani P.,
RA Romano A., Di Fiore P.P., Foiani M.;
RT "Activation of Rad53 kinase in response to DNA damage and its effect
RT in modulating phosphorylation of the lagging strand DNA polymerase.";
RL EMBO J. 18:6561-6572(1999).
RN [7]
RN STRUCTURE BY NMR OF 573-730.
RX MEDLINE=20057864; PubMed=10588905;
RA Liao H., Byeon I.-J.B., Tsai M.-D.;
RT "Structure and function of a new phosphopeptide-binding domain
RT containing the FHA2 of Rad53.";
RL J. Mol. Biol. 294:1041-1049(1999).
CC -!- FUNCTION: Controls S-phase checkpoint as well as G1 and G2 DNA
CC damage checkpoints. Phosphorylates proteins on serine, threonine,
CC and tyrosine. Prevents entry into anaphase and mitotic exit after
CC DNA damage via regulation of the Polo kinase CDC5.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- PTM: Autophosphorylated.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -!- SIMILARITY: Contains 2 FHA domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M55623; AAA35070.1; -;
DR EMBL; X96770; CAAG5568.1; -;
DR EMBL; Z73509; CAA97858.1; -;
DR FIR; A39616; A39616.
DR PDB; 1DMZ; 06-JAN-00.
DR PDB; 1FHQ; 18-OCT-00.
DR PDB; 1FHR; 18-OCT-00.
DR PDB; 1G3G; 10-JAN-01.
DR PDB; 1G6G; 13-DEC-00.
DR PDB; 1J4K; 05-DEC-01.
DR PDB; 1J4L; 05-DEC-01.
DR PDB; 1J4P; 05-DEC-01.
DR PDB; 1J4Q; 05-DEC-01.
DR PDB; 1K2M; 05-DEC-01.
DR PDB; 1K2N; 05-DEC-01.
DR PDB; 1K3J; 05-DEC-01.
DR PDB; 1K3N; 05-DEC-01.
DR PDB; 1K3Q; 05-DEC-01.
DR PDB; 1Q05; 17-JAN-01.
DR Germline; 144135; -;
DR SGD; S0006074; RAD53.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0004712; F:protein threonine/tyrosine kinase activity; IDA.
DR GO; GO:0006281; F:DNA repair; IMP.
DR GO; GO:0006139; P:nucleobase, nucleoside, nucleotide and nucl. . . ; IGI.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_kin_AS.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008984; SMAD_FHA.

DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00498; FHA; 2.
DR PRINTS; PR00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; FHA_2.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00066; FHA_DOMAIN; 2.
KW Transferase: Serine/threonine-protein kinase; ATP-binding;
KW Tyrosine-protein kinase; Nuclear protein; Repeat; DNA damage;
KW Cell cycle; Phosphorylation; 3D-structure.
FT DOMAIN 66 116 FHA 1.
FT DOMAIN 198 466 PROTEIN_KINASE.
FT NP_BIND 204 212 ATP (BY SIMILARITY).
FT BINDING 227 227 ATP (BY SIMILARITY).
FT ACT_SITE 319 319 BY SIMILARITY.
SQ SEQUENCE 821 AA; 91962 MW; 8A9612229CA72D1 CRC64;
Query Match 28.4%; Score 107.5; DB 1; Length 821;
Best Local Similarity 37.7%; Pred. No. 0.0002;
Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;
QY 5 KREWTIGRRGCDLSFFSNKLVSVDHCRIVVDEKSGQVLTEDTSTGVINKLVKKQT 64
DB 63 KKWTFGRNPAQDYLGNISLSNKHQILLGE-DGNLLNDISTGTWLNQKVEKNS 121
QY 65 CPLQTGDVVI 73
DB 122 QLLSQGDEI 130
RESULT 2
CDS1_SCHPO
ID CDS1_SCHPO STANDARD; PRT; 460 AA.
AC Q09170; O42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein kinase cds1 (EC 2.7.1.37) (Checkpoint kinase cds1).
GN CDS1 OR SPCC1835.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=95240713; PubMed=7723827;
RA Murakami H., Okayama H.;
RT "A kinase from fission yeast responsible for blocking mitosis in S
RT phase.";
RL Nature 374:817-819(1995).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND PHOSPHORYLATION.
RC STRAIN=972;
RX MEDLINE=98119835; PubMed=9450932;
RA Lindsey H.D., Griffiths D.J.F., Edwards R.J., Christensen P.U.,
RA Murray J.M., Osman F., Walworth N., Carr A.M.;
RT "S-phase-specific activation of Cds1 kinase defines a subpathway of
RT the checkpoint response in Schizosaccharomyces pombe.";
RL Genes Dev. 12:382-395(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21648401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

P16715 vaccinia vi
P33817 variola vir
P41846 caenorhabdi
P34631 caenorhabdi
Q9nsd9 homo sapien
P35172 saccharomyc
P25155 gallus gall
P39521 saccharomyc
Q9zcu5 rattus norv
Q8trs3 methanosarc
P49305 rhizobium m
Q9ru92 deinococcus

34 58 15.3 891 1 VP4A VACCV
35 58 15.3 892 1 VP4A VARV
36 58 15.3 1009 1 WS14 CAEEL
37 57.5 15.2 469 1 Y0J8 CAEEL
38 57.5 15.2 589 1 SYFB HUMAN
39 57.5 15.2 780 1 TREB YEAST
40 57 15.1 475 1 FA10 CHICK
41 57 15.1 936 1 FHL1 YEAST
42 57 15.1 1333 1 ADO RAT
43 56.5 14.9 215 1 KAD METAC
44 56.5 14.9 334 1 YNO1 RHIME
45 56.5 14.9 376 1 DNJ1 DEIRA

ALIGNMENTS

RESULT 1
ID SPK1_YEAST STANDARD; PRT; 821 AA.
AC P22216;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein kinase SPK1 (EC 2.7.1.-) (Serine-protein kinase 1).
GN SPK1 OR SADI OR RAD53 OR MEC2 OR VPL153C OR P2568.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91117267; PubMed=1899289;
RA Stern D.F., Zheng P., Beidler D.R., Zerillo C.;
RT "Spk1, a new kinase from Saccharomyces cerevisiae, phosphorylates
proteins on serine, threonine, and tyrosine";
EL Mol. Cell. Biol. 11:987-1001(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RC MEDLINE=97103777; PubMed=8948103;
RA Purnelle B., Coster F., Goffeau A.;
RT "The sequence of 55 kb on the left arm of yeast chromosome XVI
identifies a small nuclear RNA, a new putative protein kinase and two
new putative regulators";
EL Yeast 12:1483-1492(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RC MEDLINE=97133271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
Kemp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
Mueller-Auer S., Namath A., Newnith U., Oefner P., Pearson D.,
Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
Schrems B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
Urrestarazu L.A., Ushinsky S., Vierdeels F., Viessers S., Voss H.,
Walsh S.V., Wambutt R., Wang Y., Wedler E., Winnett E.,
Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI";
EL Nature 387:103-105(1997).
RN [4]
RP FUNCTION.
RX MEDLINE=95047382; PubMed=7958905;
RA Allen J.B., Zhou Z., Siede W., Friedberg E.C., Elledge S.J.;
RT "The SADI/RAD53 protein kinase controls multiple checkpoints and DNA

US-10-048-046-2_COPY_31_103
378
Sequence: 1 VLLKREWTIGRRGCDLSP.....INKLKVVKKQTCPLQTGDVI 73

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	107.5	28.4	821	1 SPK1_YEAST	P22216 saccharomyc
2	77.5	20.5	460	1 CSN1_SCHPO	Q09170 schizosacch
3	72	19.0	513	1 DUN1_YEAST	P39009 saccharomyc
4	71	18.8	409	1 PGK_MYCPN	P78018 mycoplasma
5	66	17.5	977	1 DLP1_HUMAN	O14490 homo sapien
6	66	17.5	992	1 DLP1_RAT	P97836 rattus norv
7	64	16.9	1441	1 VGLM_BUNGH	P04875 bunyavirus
8	63.5	16.8	445	1 VGLM_SCHPO	Q10292 schizosacch
9	61.5	16.3	476	1 LOK_DROME	O61267 drosophila
10	61	16.1	484	1 FKX1_YEAST	P40466 saccharomyc
11	61	16.1	546	1 CHK2_MOUSE	Q92265 mus musculus
12	61	16.1	1433	1 VGLM_BUNY7	P04505 bunyawera
13	61	16.1	1441	1 VGLM_BUNL7	P03612 bunyavirus
14	60.5	16.0	365	1 TRMU_CORGL	Q8nr24 corynebacte
15	60	15.9	353	1 Y634_CHLMU	Q9pk39 chlamydia m
16	60	15.9	403	1 G6ST_MOUSE	Q8cfa2 mus musculus
17	60	15.9	952	1 YK15_CAEEL	P45012 caenorhabdi
18	60	15.9	969	1 UVRA_MYLE	Q9cc24 mycobacteri
19	59.5	15.7	465	1 TM15_HUMAN	Q9c019 homo sapien
20	59.5	15.7	485	1 RNFB_HUMAN	O76064 homo sapien
21	59.5	15.7	809	1 DCM1_OLICA	P19919 oligotropho
22	59	15.6	592	1 YTRP_PSEPU	P40604 pseudomonas
23	59	15.6	1045	1 HMD2_YEAST	P12684 saccharomyc
24	58.5	15.5	181	1 KAD_WCTUT	O53796 mycobacteri
25	58.5	15.5	209	1 GPNB_RHET	Q8kl44 rhizobium e
26	58.5	15.5	523	1 P60_LISSE	Q01838 listeria se
27	58.5	15.5	814	1 SLAI_BACAA	P49051 bacillus an
28	58	15.3	497	1 MEK1_YEAST	P24719 saccharomyc
29	58	15.3	522	1 YNL6_YEAST	P53924 saccharomyc
30	58	15.3	524	1 HUNB_TRICA	Q01791 tribolium c
31	58	15.3	746	1 TAGP_BACSU	P13485 bacillus su
32	58	15.3	888	1 SYA_ZYMO	Q9rn98 zymomonas m
33	58	15.3	891	1 VP4A_VACCC	P20642 vaccinia vi

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 20:06:30 ; Search time 1483 Seconds

(without alignments)
7674.257 Million cell updates/sec

Title: US-10-048-046-1

Perfect score: 2679

Sequence: 1 aagaaatcgacagagggccg.....acaaaaa..... 2679

Scoring table: OLIGO NUC

Gapop 60.0, Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 29Jan04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2679	100.0	2679	5 AAF30352	Human chf
2	2133	79.6	2639	4 AAF14542	Human CDN
3	1995	74.5	1995	5 AAF89709	Nucleotid
4	1618	60.4	3181	6 AAF59646	Novel hum
5	1465	54.7	3138	4 AAF14556	Human CDN
6	1024	38.2	2259	4 AAF188903	Human pol
7	765	28.6	1311	4 AAF25843	Human CDN
8	765	28.6	1311	7 AAF73184	Human nov
9	633	23.6	2186	7 ADA52592	Human cod
10	478	17.8	816	4 AAF06828	Human CDN
11	468	17.5	824	4 AAF06846	Human CDN
12	455	17.0	693	4 AAF26304	Human nov
13	455	17.0	693	7 AAF73645	Human CDN
14	426	15.9	518	4 AAF11859	Human sec
15	286	10.7	575	3 AAF44336	Human CDN
16	267	10.0	357	4 AAF26143	Human nov
17	267	10.0	357	7 AAF73484	Human CDN
18	238	8.9	354	4 AAF26563	Human CDN
19	238	8.9	354	4 AAF26144	Human CDN
20	238	8.9	354	7 AAF73904	Human nov
21	238	8.9	354	7 AAF73485	Human nov
22	217	8.1	449	8 ACH26450	Human adu
23	169	6.3	13836	5 ABA16633	Human ner

ALIGNMENTS

RESULT 1

AAF30352
ID AAF30352 standard; cDNA; 2679 BP.

XX AAF30352;

DT 14-MAY-2001 (first entry)

XX Human chfr cDNA encoding checkpoint with FHA and ring finger protein.

DE Checkpoint with forkhead associated domain and ring finger; Chfr; human;

KW Mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening;

KW ubiquitin-protein ligase; ss.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 91..2085

FT /*tag= a

FT mutation replace(1828,A)

FT /*tag= b

FT /note= "alters codon GTG (Val) to ATG (Met) in chfr from

human cancer cell line U2OS"

FT WO200109150-A2.

XX 08-FEB-2001.

XX 14-JUN-2000; 2000WO-US016391.

XX 29-JUL-1999; 99US-0146194P.

XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX Halazonetis T, Scolnick D;

XX WPI; 2001-182927/18.

XX P-PSDB; AAB20219.

XX Novel nucleic acid sequence of mitotic checkpoint gene encoding a

XX checkpoint with forkhead-associated domain and ring finger protein, for

XX diagnosing tumorigenic cells and in screening for anticancer drugs.

XX Claim 2(a); Fig 4A-D; 85pp; English.

XX The present sequence is that of cDNA encoding the human mitotic

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

checkpoint protein Chfr (see AAB20219) having a forkhead associated domain (FHA) and a ring finger domain. The protein is required for regulation of the transition of cells from prophase to metaphase during mitosis. It has ubiquitin-protein ligase activity. The Chfr checkpoint is evident in primary human cells, but was inactivated in 4 of 8 human cancer cell lines. In U2OS cells, a mutation was identified, which caused a Val to Met amino acid substitution in the highly conserved C-terminal Cys-rich region of the Chfr protein. In the absence of the Chfr checkpoint, cells subjected to mitotic stress condensed their chromosomes despite failing to separate their chromosomes. Chfr may monitor centrosome separation. Inactivation of the Chfr gene in human cancer is theorized to underlie the increased sensitivity of cancer cells to antimitotic drugs. Nucleic acids comprising the present sequence, or sequences encoding at least amino acids 31-103, 303-346, 476-641 (or their antisense sequences) are claimed. The Chfr cDNA was isolated from an expressed sequence tag database sequence for cDNAs with FHA motifs. Claimed methods of determining the tumorigenic potential of a cell comprise examining the cell for the presence of Chfr nucleic acid (absence or mutation indicating predisposition to tumorigenesis upon exposure to mitotic stress). Also claimed is a diagnostic agent, e.g., an antisense fragment of the present sequence, that binds to the Chfr nucleic acid, and a diagnostic kit for detecting tumorigenic potential of a cell. A composition which inhibits the biological activity of Chfr may comprise a ligand selected from an antibody or its fragment. The Chfr inhibitor is used in a claimed method of retarding the growth of a cancer cells. Chfr polypeptides are useful in screening for drugs which can inhibit the activity of Chfr in a cancer cell, rendering the cell more sensitive to additional antitumor therapies

Sequence 2679 BP; 650 A; 715 C; 764 G; 550 T; 0 U; 0 Other;

Query Match 100.0%; Score 2679; DB 5; Length 2679;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGATTCCGACAGCGCGCGGATCTCTTGACAGCGCGCGCGCGCGCGGTTCCGG 60
 DB 1 AAGATTCCGACAGCGCGCGGATCTCTTGACAGCGCGCGCGCGCGCGGTTCCGG 60
 QY 61 GTTCGCGCGCGCGCGGATGTGAATCCCGATGAGCGCGCGCGCGCGCGGCAAGCAGTCG 120
 DB 61 GTTCGCGCGCGCGCGGATGTGAATCCCGATGAGCGCGCGCGCGCGGCAAGCAGTCG 120
 QY 121 CGCGCGCGCGCGCGGATGTGAATCCCGATGAGCGCGCGCGCGCGGCAAGCAGTCG 180
 DB 121 CGCGCGCGCGCGCGGATGTGAATCCCGATGAGCGCGCGCGCGCGGCAAGCAGTCG 180
 QY 181 GTTCCTCTGAGGAAGCGGAGTGAGACCATCGCGCGCGCGCGCGGCAAGCAGTCG 240
 DB 181 GTTCCTCTGAGGAAGCGGAGTGAGACCATCGCGCGCGCGCGCGGCAAGCAGTCG 240
 QY 241 CCCAGCATTAATCTGCTCTGAGATCACTGTAGATTTAGTGGATGAATAATCAGGT 300
 DB 241 CCCAGCATTAATCTGCTCTGAGATCACTGTAGATTTAGTGGATGAATAATCAGGT 300
 QY 301 CAGGTGACACTGGAAGATACCAAGCAGGATGATTAACAAAGCTGAAGTTGTT 360
 DB 301 CAGGTGACACTGGAAGATACCAAGCAGGATGATTAACAAAGCTGAAGTTGTT 360
 QY 361 AAGAGCAGACATGCTCTTACAGATCGGATGATCTACTTGTGTACAGAGAAAT 420
 DB 361 AAGAGCAGACATGCTCTTACAGATCGGATGATCTACTTGTGTACAGAGAAAT 420
 QY 421 GAACCGGAACACAGCTGGGATACCTCTATGAATCTTTAAGTGAAGGCAAGGATGACA 480
 DB 421 GAACCGGAACACAGCTGGGATACCTCTATGAATCTTTAAGTGAAGGCAAGGATGACA 480
 QY 481 CAGAGATCTTTGAGCTTAAAGAGAAATCTGTTCATGGACCAAGATACCTCAGGT 540
 DB 481 CAGAGATCTTTGAGCTTAAAGAGAAATCTGTTCATGGACCAAGATACCTCAGGT 540
 QY 541 GCAGGTGACAGGCGGCGGATCCCGGGTCCCTCGCTCGCGCGCGGCACTCAGGTG 600
 DB 541 GCAGGTGACAGGCGGCGGATCCCGGGTCCCTCGCTCGCGCGGCACTCAGGTG 600

DB 541 GCAGGTGACAGGCGGCGGATCCCGGGTCCCTCGCTCGCGCGGCACTCAGGTG 600
 QY 601 TCGTTTGGAGAACACACAGCATCAACATCGAGTCAAGCTTTCCCAACAGCCTCGGCC 660
 DB 601 TCGTTTGGAGAACACACAGCATCAACATCGAGTCAAGCTTTCCCAACAGCCTCGGCC 660
 QY 661 TCTTCCACAGGAGCCTTCTCTGACAGGCGGAGCGTTCCTCCAGTCTGGGTGCGGGT 720
 DB 661 TCTTCCACAGGAGCCTTCTCTGACAGGCGGAGCGTTCCTCCAGTCTGGGTGCGGGT 720
 QY 721 GTTGCGATCTTCCCTAAAGAGTGGTCCCTCTCTGTGGGAAGTGAAGTCTCCAGCTTT 780
 DB 721 GTTGCGATCTTCCCTAAAGAGTGGTCCCTCTCTGTGGGAAGTGAAGTCTCCAGCTTT 780
 QY 781 GCGTCAGTCTTCCCAACAGCAAGAACTGCGCTCTTTTCTGTGGTGAACCCAGGATCAG 840
 DB 781 GCGTCAGTCTTCCCAACAGCAAGAACTGCGCTCTTTTCTGTGGTGAACCCAGGATCAG 840
 QY 841 GAGGATTTGGAGCCCGTGAAGAAATGAGAGGATGGGACCTTGACCTGAAACGGG 900
 DB 841 GAGGATTTGGAGCCCGTGAAGAAATGAGAGGATGGGACCTTGACCTGAAACGGG 900
 QY 901 CAGTTTGTGGTGGACAAACCCGTAGAAATGCCCAAACCGTCCACAGGAGCGTCAGAGCA 960
 DB 901 CAGTTTGTGGTGGACAAACCCGTAGAAATGCCCAAACCGTCCACAGGAGCGTCAGAGCA 960
 QY 961 GCGGCTGGAGAACACAGCAAGATGAGAGGAGCGTGAATCATCTGCCAGGACCTG 1020
 DB 961 GCGGCTGGAGAACACAGCAAGATGAGAGGAGCGTGAATCATCTGCCAGGACCTG 1020
 QY 1021 CTGACAGCACTCGGTGAGTTTTCAGCCCTGCAATGCACACAGTTTTCGCGCGGTTGCTAC 1080
 DB 1021 CTGACAGCACTCGGTGAGTTTTCAGCCCTGCAATGCACACAGTTTTCGCGCGGTTGCTAC 1080
 QY 1081 GCGTGGATGAGAGCGCTGCTCTGCTGCTTACCTGCGCGCTGCTCCGCTGGAGCGGATCTGT 1140
 DB 1081 GCGTGGATGAGAGCGCTGCTCTGCTGCTTACCTGCGCGCTGCTCCGCTGGAGCGGATCTGT 1140
 QY 1141 AAAAACCACATCTCTCAACCACTCGTGAAGCAATCCTCATCCAGCATCCAGCAAGAGT 1200
 DB 1141 AAAAACCACATCTCTCAACCACTCGTGAAGCAATCCTCATCCAGCATCCAGCAAGAGT 1200
 QY 1201 CGCATGAGAGGATGCGAAGTATGATGCCAGGATTAATCACTCAAGACATGCTG 1260
 DB 1201 CGCATGAGAGGATGCGAAGTATGATGCCAGGATTAATCACTCAAGACATGCTG 1260
 QY 1261 CAGCCCAAGTCAGCGCGCTCTTTTCTGATGAAGAGGAGTTTCAGAGGACCTGCTGGAG 1320
 DB 1261 CAGCCCAAGTCAGCGCGCTCTTTTCTGATGAAGAGGAGTTTCAGAGGACCTGCTGGAG 1320
 QY 1321 CTGTCAAGCTTGACAGTGTCTCTGACATTTAGCCAGCATACCTGCTGTCGCGGAG 1380
 DB 1321 CTGTCAAGCTTGACAGTGTCTCTGACATTTAGCCAGCATACCTGCTGTCGCGGAG 1380
 QY 1381 TGTCTGAGTACAGAGGCGGCGGCGAGCTTCCCACTGCGCCAGCAACCCAGGCGGAG 1440
 DB 1381 TGTCTGAGTACAGAGGCGGCGGCGAGCTTCCCACTGCGCCAGCAACCCAGGCGGAG 1440
 QY 1441 CCAGAGCCCGCACAGGCGCTGGGGGATGACCCCTCCAGCTCCCTGAGCTGACAGCA 1500
 DB 1441 CCAGAGCCCGCACAGGCGCTGGGGGATGACCCCTCCAGCTCCCTGAGCTGACAGCA 1500
 QY 1501 GTCCAGGATTAACGTGTGCTTCTGCAAGGAAGCAAGCCCTGTCACCTGCTTCCAG 1560
 DB 1501 GTCCAGGATTAACGTGTGCTTCTGCAAGGAAGCAAGCCCTGTCACCTGCTTCCAG 1560
 QY 1561 CCCATGCGCGACCGGAGCGGAGCGGAGCGCGGCTGCTCCCTCAGCAGGTGT 1620
 DB 1561 CCCATGCGCGACCGGAGCGGAGCGGAGCGGCTGCTCCCTCAGCAGGTGT 1620
 QY 1621 CCGGTCTCTGACAGCCTTTCTGCACTTGTACTGGGGTGTGACCCCGGACCGGCTGCTAC 1680
 DB 1621 CCGGTCTCTGACAGCCTTTCTGCACTTGTACTGGGGTGTGACCCCGGACCGGCTGCTAC 1680

QY 1681 GGCTGCTGGCCCGCTTTTGTGAGCTCAACCTGGGTGACAGTGTCTGGAAGCGCTGCTG 1740
Db 1681 GGCTGCTGGCCCGCTTTTGTGAGCTCAACCTGGGTGACAGTGTCTGGAAGCGCTGCTG 1740
QY 1741 AACACAAACAGCTACAGGTGACAGATCTCTGAAGATTTACCTGGCAACAGAGTTTGACA 1800
Db 1741 AACACAAACAGCTACAGGTGACAGATCTCTGAAGATTTACCTGGCAACAGAGTTTGACA 1800
QY 1801 TGGAAAAACATGTTGACCGAGAGCTCTGGGTCTCTCCAGCGGGAGTGTCTTCTGTCT 1860
Db 1801 TGGAAAAACATGTTGACCGAGAGCTCTGGGTCTCTCCAGCGGGAGTGTCTTCTGTCT 1860
QY 1861 GATTACAGAGTCAACGGAGACACCGTCTGTGTGTTACTGTGTGGCTTGGCGAGCTTCCGT 1920
Db 1861 GATTACAGAGTCAACGGAGACACCGTCTGTGTGTTACTGTGTGGCTTGGCGAGCTTCCGT 1920
QY 1921 GAGCTGACCTATCAGTATCGGAGACATCTCTGCTTCGGAGTGTGCAAGTGGCCGTAAACA 1980
Db 1921 GAGCTGACCTATCAGTATCGGAGACATCTCTGCTTCGGAGTGTGCAAGTGGCCGTAAACA 1980
QY 1981 TCCCGTCTGACTGCTACTGCGGCCGCTAACTGCGCACTCAGGTGAAGCTCACCAGCC 2040
Db 1981 TCCCGTCTGACTGCTACTGCGGCCGCTAACTGCGCACTCAGGTGAAGCTCACCAGCC 2040
QY 2041 ATGAATTCATATCATATCTGTGACAGACAGGTTCAAAATCTAAGCATCCAGAGGCCCT 2100
Db 2041 ATGAATTCATATCATATCTGTGACAGACAGGTTCAAAATCTAAGCATCCAGAGGCCCT 2100
QY 2101 GAGCAGCTTTTCAAGCTGAGAGTGAAGAGAGCGTGTGTTTTAAATAACAGAGCAAGCAG 2160
Db 2101 GAGCAGCTTTTCAAGCTGAGAGTGAAGAGAGCGTGTGTTTTAAATAACAGAGCAAGCAG 2160
QY 2161 TCAAGTGTGTTTCAAGCGCCCTGAGGAGGAGGAGCGAGGCTCTCGACAGTGTCTGCG 2220
Db 2161 TCAAGTGTGTTTCAAGCGCCCTGAGGAGGAGGAGCGAGGCTCTCGACAGTGTCTGCG 2220
QY 2221 GGTGACTCTTCTGTGAGCTTTTACCTCTGAGTGAGACCTCTCCAGAGCCCGCGGG 2280
Db 2221 GGTGACTCTTCTGTGAGCTTTTACCTCTGAGTGAGACCTCTCCAGAGCCCGCGGG 2280
QY 2281 CCGCAGCCCGCCCTCTGAGCGCTGAGCGGCTCTGTTGGCATCAGCAGCAGAGA 2340
Db 2281 CCGCAGCCCGCCCTCTGAGCGCTGAGCGGCTCTGTTGGCATCAGCAGCAGAGA 2340
QY 2341 CGAAGCTTTTCTTAACATGCGGCCCTCCCGCAGAGGGGCGAGTTTGTCTTTTGTAC 2400
Db 2341 CGAAGCTTTTCTTAACATGCGGCCCTCCCGCAGAGGGGCGAGTTTGTCTTTTGTAC 2400
QY 2401 ATTTTCGAAACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAGTTTCAAGGGGAGAG 2460
Db 2401 ATTTTCGAAACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAGTTTCAAGGGGAGAG 2460
QY 2461 GGCAAGTTTATCAAAACATGTTTTCAGGAGAGGGGAGATGTTTACAGCTCAGAGA 2520
Db 2461 GGCAAGTTTATCAAAACATGTTTTCAGGAGAGGGGAGATGTTTACAGCTCAGAGA 2520
QY 2521 CGTACAAATATCTGCTGCTGGGAAACACACAGCATTTTATCTATTTTATTTAATA 2580
Db 2521 CGTACAAATATCTGCTGCTGGGAAACACACAGCATTTTATCTATTTTATTTAATA 2580
QY 2581 GGTGTTGCTTATCTTCTTAATTAAGATTTTAAATGTCAAACTGTAGCAAAATATATA 2640
Db 2581 GGTGTTGCTTATCTTCTTAATTAAGATTTTAAATGTCAAACTGTAGCAAAATATATA 2640
QY 2641 ATTTTATTAATTTACAAATTTGACAAAAAATTTTAAAAA 2679
Db 2641 ATTTTATTAATTTACAAATTTGACAAAAAATTTTAAAAA 2679

RESULT 2
AAH14542
ID AAH14542 standard; cDNA; 2639 BP.

XX AAH14542;
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:12099.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX Claim 8; SEQ ID NO 12099; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX Sequence 2639 BP; 620 A; 706 C; 760 G; 553 T; 0 U; 0 Other;
Query Match 79.6%; Score 2133; DB 4; Length 2639;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 529 GATACCTCAGGTGCGAGGTGCGAGGGCGGCGGATCCCGGGTCCCTCGTCTGCGCC 588
Db 481 GATACCTCAGGTGCGAGGTGCGAGGGCGGCGGATCCCGGGTCCCTCGTCTGCGCC 540
QY 589 GCCACTCAGGTGCTGTTTGGAGAACCCAGCCATCAGCATCAGCTCAGACTCTTCCCC 648
Db 541 GCCACTCAGGTGCTGTTTGGAGAACCCAGCCATCAGCATCAGCTCAGACTCTTCCCC 600

QY 649 ACAGCTCGGCTCTTCCACGAGGCTTCTCCTCAGGCGGAGAGCCTTCTCCTCAGTTGT 708
Db 601 ACAGCTCGGCTCTTCCACGAGGCTTCTCCTCAGGCGGAGAGCCTTCTCCTCAGTTGT 660
QY 709 GGGTCTGGGGTGTGGGATCTCCCTTAAGGAAAGTGTCTCCTCTGCGCAAGTGATGAA 768
Db 661 GGGTCTGGGGTGTGGGATCTCCCTTAAGGAAAGTGTCTCCTCTGCGCAAGTGATGAA 720
QY 769 GTCTCCAGCTTTGGCTCAGCTCTCCACAGCAAGAGCTCGCTCCTTTTCTCGTTGGAA 828
Db 721 GTCTCCAGCTTTGGCTCAGCTCTCCACAGCAAGAGCTCGCTCCTTTTCTCGTTGGAA 780
QY 829 CCCAGGATCAGGAGGATTTGGAGCCCTGAGAGAAATGAGAGAGATGGGAGCCTT 888
Db 781 CCCAGGATCAGGAGGATTTGGAGCCCTGAGAGAAATGAGAGAGATGGGAGCCTT 840
QY 889 GACCTGAACGGGCACTGTCTGTGCGCAACCGGCTGAGAAATGCCAAACCGTCCACGAG 948
Db 841 GACCTGAACGGGCACTGTCTGTGCGCAACCGGCTGAGAAATGCCAAACCGTCCACGAG 900
QY 949 GACGTACAGCAGCGCTGGAGAGCCAGCAAGATGAGAGAGAGCTGACATGATCATC 1008
Db 901 GACGTACAGCAGCGCTGGAGAGCCAGCAAGATGAGAGAGAGCTGACATGATCATC 960
QY 1009 TGCCAGGACCTGCTGCACAGACTGCGTGAGTTTGAGCCCTGCATGCAACAGCTTCTGGCGG 1068
Db 961 TGCCAGGACCTGCTGCACAGACTGCGTGAGTTTGAGCCCTGCATGCAACAGCTTCTGGCGG 1020
QY 1069 GCTTCTACTCGGGCTGAGATGAGCGCTGCTCCTCTGTGCTCTACTGCGCGCTGTCCGCTG 1128
Db 1021 GCTTCTACTCGGGCTGAGATGAGCGCTGCTCCTCTGTGCTCTACTGCGCGCTGTCCGCTG 1080
QY 1129 GAGCGGATCTGTAAGAACCATCTCTCAACACCTCTGTAAGCATACCTCATCCAGCAT 1188
Db 1081 GAGCGGATCTGTAAGAACCATCTCTCAACACCTCTGTAAGCATACCTCATCCAGCAT 1140
QY 1189 CCAGCAAGAGTCGAGTGAAGAGATGCAAGATATGATGATGCAAGAAATAAATCACT 1248
Db 1141 CCAGCAAGAGTCGAGTGAAGAGATGCAAGATATGATGATGCAAGAAATAAATCACT 1200
QY 1249 CAAGCATGCTGCAGCCCAAGTCAAGGCTGCTTTCTGATGAGAGGAGGATTCAGAG 1308
Db 1201 CAAGCATGCTGCAGCCCAAGTCAAGGCTGCTTTCTGATGAGAGGAGGATTCAGAG 1260
QY 1309 GACCTGCTGGAGCTGTGACAGCTTGACAGTGTCTCTCAGACATTTAGCCAGCCATAGCTC 1368
Db 1261 GACCTGCTGGAGCTGTGACAGCTTGACAGTGTCTCTCAGACATTTAGCCAGCCATAGCTC 1320
QY 1369 GTGTGCGGCACTGCTGATGATACAGAGGAGGCGGCGAGCCTCCCACTGCGCCAGCA 1428
Db 1321 GTGTGCGGCACTGCTGATGATACAGAGGAGGCGGCGAGCCTCCCACTGCGCCAGCA 1380
QY 1429 CCGAGGCGAGCCAGAGCCCAAGGCTCTGGGGGATGCACCTCCACGCTCCGTCCAGC 1488
Db 1381 CCGAGGCGAGCCAGAGCCCAAGGCTCTGGGGGATGCACCTCCACGCTCCGTCCAGC 1440
QY 1489 CTGACACAGCAGTCCAGGATACGTTGTCCTCTGCAAGGAGCCAGCCCTGTGACCC 1548
Db 1441 CTGACACAGCAGTCCAGGATACGTTGTCCTCTGCAAGGAGCCAGCCCTGTGACCC 1500
QY 1549 TGTCTGCTCCAGCCATGCCGACGAGCGAGCGGAGCAGACCCGCGTGTGCGC 1608
Db 1501 TGTCTGCTCCAGCCATGCCGACGAGCGGAGCGGAGCAGACCCGCGTGTGCGC 1560
QY 1609 CTTACAGCAGTGTGGGTCTGCTGCAAGCTTTCTGCCACTTACTTGGGGCTGCAACCGG 1668
Db 1561 CTTACAGCAGTGTGGGTCTGCTGCAAGCTTTCTGCCACTTACTTGGGGCTGCAACCGG 1620
QY 1669 ACCGGTGTACGGCTGCTGCGTGTGCGTGTGAGCTCAACCTGGGTGACAGTGTCTG 1728
Db 1621 ACCGGTGTACGGCTGCTGCGTGTGCGTGTGAGCTCAACCTGGGTGACAGTGTCTG 1680

QY 1729 GACGCGTCTGAAACAACAAGCTACGAGTCAGATCCCTGAAGAAATTAACCTGCAACC 1788
Db 1681 GACGCGTCTGAAACAACAAGCTACGAGTCAGATCCCTGAAGAAATTAACCTGCAACC 1740
QY 1789 AGAGGTTTGACATGAGAAACAAATGTTGACCGAGAGCCCTGCTGCTCCAGCGGGAGTG 1848
Db 1741 AGAGGTTTGACATGAGAAACAAATGTTGACCGAGAGCCCTGCTGCTCCAGCGGGAGTG 1800
QY 1849 TTTCTGCTGCTGATTTACAGAGTCACGGGAGACACCTGTTCTGTTTACTGCTGTGGCCTG 1908
Db 1801 TTTCTGCTGCTGATTTACAGAGTCACGGGAGACACCTGTTCTGTTTACTGCTGTGGCCTG 1860
QY 1909 CGCAGCTTCCGTGAGCTGACCTTATCAGTATCCGCAAGAACATTCCTGCTTCCGAGTTGCCA 1968
Db 1861 CGCAGCTTCCGTGAGCTGACCTTATCAGTATCCGCAAGAACATTCCTGCTTCCGAGTTGCCA 1920
QY 1969 GTGGCCGTAACATCCCGTCTGACTGCTACTGGGCGCTTAACCTGCGCACTCAGGTGAAA 2028
Db 1921 GTGGCCGTAACATCCCGTCTGACTGCTACTGGGCGCTTAACCTGCGCACTCAGGTGAAA 1980
QY 2029 GCTCACACGCGCATGAAATTCATATCTGTGAACAGCAAGGTTCAAAAACCTAAGCA 2088
Db 1981 GCTCACACGCGCATGAAATTCATATCTGTGAACAGCAAGGTTCAAAAACCTAAGCA 2040
QY 2089 TCCAGAGGCGCTGAGCAGCTTTTACGACTGAGAGTGAAGAGCCGCTTTTAAATACA 2148
Db 2041 TCCAGAGGCGCTGAGCAGCTTTTACGACTGAGAGTGAAGAGCCGCTTTTAAATACA 2100
QY 2149 GAGCAAGCAGCTCAAGTGTTTTCACAGCCCTTGAAGGAAAGGAGCGCAGGCTTCCGA 2208
Db 2101 GAGCAAGCAGCTCAAGTGTTTTCACAGCCCTTGAAGGAAAGGAGCGCAGGCTTCCGA 2160
QY 2209 CAGGTCTCTGGGCTGACTCTCTCTGTGAGCTTTTACCTCTGAGTGAGACCTCCCA 2268
Db 2161 CAGGTCTCTGGGCTGACTCTCTCTGTGAGCTTTTACCTCTGAGTGAGACCTCCCA 2220
QY 2269 GAGCCCGGGCGCGCAGCCCTCTGCTGAGCGCTGCGCAGGCTCGTGTGGCAT 2328
Db 2221 GAGCCCGGGCGCGCAGCCCTCTGCTGAGCGCTGCGCAGGCTCGTGTGGCAT 2280
QY 2329 CAGCAGCAGAGAGCAAGCCTTTCTGTAAATGCGGCGCTCCCGCCGAGAGGGGCGAGTTT 2388
Db 2281 CAGCAGCAGAGAGCAAGCCTTTCTGTAAATGCGGCGCTCCCGCCGAGAGGGGCGAGTTT 2340
QY 2389 GCTCTTTTGTACTTTTCCGAACTACAGTTAAAGCAGAACTGTGTTTTCAGGAAAGTT 2448
Db 2341 GCTCTTTTGTACTTTTCCGAACTACAGTTAAAGCAGAACTGTGTTTTCAGGAAAGTT 2400
QY 2449 TCAAGGAGAGAGGCGCAAGTTTATCAAAAACATTTTCAAGGAGAGGAGCATAGTTTA 2508
Db 2401 TCAAGGAGAGAGGCGCAAGTTTATCAAAAACATTTTCAAGGAGAGGAGCATAGTTTA 2460
QY 2509 CAGCCTACAGAGCTACACAAATCTCTGCTGCTGGGAAACCAACAGCATTTTATCTATT 2568
Db 2461 CAGCCTACAGAGCTACACAAATCTCTGCTGCTGGGAAACCAACAGCATTTTATCTATT 2520
QY 2569 TTTATTTTAAATAGGTTTGTGCTTATCTCTTAATAAGATTTTAAATCTCAAACTGTAGC 2628
Db 2521 TTTATTTTAAATAGGTTTGTGCTTATCTCTTAATAAGATTTTAAATCTCAAACTGTAGC 2580
QY 2629 ACAAAATAATAATTTATTAATTTTACAAATGAC 2661
Db 2581 ACAAAATAATAATTTATTAATTTTACAAATGAC 2613

RESULT 3
AAF89709
ID AAF89709 standard; DNA; 1995 BP.
XX
AC AAF89709;
XX
DT 22-AUG-2001 (first entry)
XX

DE	Nucleotide sequence of a human ring finger protein designated FHAR1.
XX	FHAR1; RING finger protein; cancer; vaccine; ss.
XX	Homo sapiens.
XX	Key Location/Qualifiers
FT	1..1195
FT	/*tag= a
FT	/product= "ring finger protein FHAR1"
XX	WO200142430-A1.
XX	14-JUN-2001.
XX	07-DEC-2000; 2000WO-US033094.
XX	08-DEC-1999; 99US-00456876.
XX	(SMIK) SMITHKLINE BEECHAM CORP.
XX	Zhou B, Zhu Y, Chaturvedi P, Hurle MR, Li X;
XX	WPI; 2001-381663/40.
XX	P-PSDB; AAB83843.
XX	New FHAR1 polypeptide, a member of the RING finger protein family for
PT	diagnosing and treating cancer, and for use in anti-cancer vaccines.
XX	Claim 2; Page 18-19; 28pp; English.
XX	The present sequence encodes a FHAR1 polypeptide, which is a member of
CC	the RING finger protein family. FHAR1 is useful in the treatment of
CC	cancer, and as a vaccine for inducing an immunological response in a
CC	mammal. FHAR1 polynucleotides may also be used as a diagnostic reagent
CC	through detection of mutations in the associated gene, and for chromosome
CC	localization studies, and tissue expression studies. FHAR1 antibodies are
CC	useful to isolate and to identify clones expressing the polypeptides, or
CC	to purify the polypeptides by affinity chromatography and to treat cancer
XX	Sequence 1995 BP; 465 A; 563 C; 580 G; 387 T; 0 U; 0 Other;
XX	Query Match 74.5%; Score 1995; DB 5; Length 1995;
XX	Best Local Similarity 100.0%; Pred. No. 0;
XX	Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	91 ATGGAGCGGCCGAGGAGGCGAGCGAGCGCGCGCGCGAGCGCTGGGACGGCTCTCG 150
Db	1 ATGGAGCGGCCGAGGAGGCGAGCGAGCGCGCGCGCGAGCGCTGGGACGGCTCTCG 60
Qy	151 CGTCTCGGCGCGAGGAGGCGAGCGCGAGCGCGAGCGCTCTCGAGGAGCGGAGTGGACATC 210
Db	61 CGTCTCGGCGCGAGGAGGCGAGCGCGAGCGCGAGCGCTCTCGAGGAGCGGAGTGGACATC 120
Qy	211 GGGCGAGGACGAGTTGCGACCTTCTTCCCGAGCGAATAAAGTGTCTCTGGAGATCAC 270
Db	121 GGGCGAGGACGAGTTGCGACCTTCTTCCCGAGCGAATAAAGTGTCTCTGGAGATCAC 180
Qy	271 TGTAGAAATTGTAGTGGATGAAAATTCAGGTGAGTGCACATGGAAGATACACGACCGAGT 330
Db	181 TGTAGAAATTGTAGTGGATGAAAATTCAGGTGAGTGCACATGGAAGATACACGACCGAGT 240
Qy	331 GGAAACAGTGATTAAACAGCTGAAGTGTGTTAAGAGCGAGACATGCCCTTTACAGACTGGG 390
Db	241 GGAAACAGTGATTAAACAGCTGAAGTGTGTTAAGAGCGAGACATGCCCTTTACAGACTGGG 300
Qy	391 GATGTCATCTACTTGTGTGTACAGGAAGATGAACCGGAACACCAACGTGGCATACCTCTAT 450
Db	301 GATGTCATCTACTTGTGTGTACAGGAAGATGAACCGGAACACCAACGTGGCATACCTCTAT 360
Qy	451 GAATCTTTAAGTGAAAGCAAGGCGATGACACAGAATCTTTTGAAGCTAAACAGGAAAT 510
Db	361 GAATCTTTAAGTGAAAGCAAGGCGATGACACAGAATCTTTTGAAGCTAAACAGGAAAT 420

Dn	2449	TTTATTTTAAGGTGGTGCTTACTTCCTAATAAGATTAAATGTCACAAACTGAGC	2500
Oy	2629	ACAATAATAATTTATATAATTTTACAAATGAC	2661
Dn	2509	ACAATAATAATTTATATAATTTTACAAATGAC	2541
RESULT 5			
ID	AHH14556		
XX	AAH14556 standard; cDNA; 3138 BP.		
AC	AAH14556;		
DT	26-JUN-2001 (first entry)		
XX	Human cDNA sequence SEQ ID NO:12127.		
DE	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
XX	Homo sapiens.		
OS	EP1074617-A2.		
PN	07-FEB-2001.		
XX	28-JUL-2000; 2000EP-00116126.		
PD	29-JUL-1999; 99JP-00248036.		
XX	27-AUG-1999; 99JP-00300253.		
PR	11-JAN-2000; 2000JP-0018776.		
PR	02-MAY-2000; 2000JP-00183767.		
PR	09-JUN-2000; 2000JP-00241899.		
XX	(HELI-) HELIX RES INST.		
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
XX	WPI; 2001-318749/34.		
DR	Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.		
XX	Claim 8; SEQ ID NO 12127; 2537pp + Sequence Listing; English.		
CC	The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:		
CC	(a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of oligonucleotides comprises at least 15 nucleotides selected from those defined in the 5'-end sequence/3'-end sequence is selected from those defined in the gene therapy. The primer sets can be used in antisense therapy and in particular full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention		
XX	Sequence 3138 BP; 738 A; 824 C; 873 G; 703 T; 0 U; 0 Other;		
SQ	Query Match 54.7%; Score 1465; DB4; Length 3138;		

Query Match

Best Local Similarity 99.6%; Pred. No. 0;			
Matches 1945; Conservative 0; Mismatches 6; Indels 2; Gaps 1;			
QY	709	GGGCTGGGGTGGTGGCATCTCCCTTAAGGAAGTGGTCCCTCTGTGGCAAGTATGAA	768
Db	560	GGGCTGGGGTGGTGGCATCTCCCTTAAGGAAGTGGTCCCTCTGTGGCAAGTATGAA	619
QY	769	GTCTCCAGCTTTGGCTCAGCTCTCCAGACAGAAAGACTGGTCCCTTTTGGTCTGGAA	828
Db	620	GTCTCCAGCTTTGGCTCAGCTCTCCAGACAGAAAGACTGGTCCCTTTTGGTCTGGAA	679
QY	829	CCCCAGATCAGGAGATTTGGAGCCGTGAAGAAAGAAATGAGAGAGATGGGACCTT	888
Db	680	CCCCAGATCAGGAGATTTGGAGCCGTGAAGAAAGAAATGAGAGAGATGGGACCTT	739
QY	889	GACCTGAAGGAGCAGTGTGTGTGCACAAACCGGTGAGAAATGCCAAACCGTCCACGAG	948
Db	740	GACCTGAAGGAGCAGTGTGTGTGCACAAACCGGTGAGAAATGCCAAACCGTCCACGAG	799
QY	949	GACGTTCAGAGCAGCGCTGGGAAGCCAGCAAGATGGAGGAGAGCGTGAATGATCATC	1008
Db	800	GACGTTCAGAGCAGCGCTGGGAAGCCAGCAAGATGGAGGAGAGCGTGAATGATCATC	859
QY	1009	TCCAGAGCCTGCTGCAGCATGCGTGAATTTGACGCGCTGCATGCAACCGTTCGCGCG	1068
Db	860	TCCAGAGCCTGCTGCAGCATGCGTGAATTTGACGCGCTGCATGCAACCGTTCGCGCG	919
QY	1069	GCTTGGCTACTCGGGTGGATGAGCGCTCGTCCCTGTGTCTTACTGCGCGTGTCCCGTG	1128
Db	920	GCTTGGCTACTCGGGTGGATGAGCGCTCGTCCCTGTGTCTTACTGCGCGTGTCCCGTG	979
QY	1129	GAGCGGATCTGTAAAAACCAATCTCAACACCTCTGGAAGCATCTCATCCAGCAT	1188
Db	980	GAGCGGATCTGTAAAAACCAATCTCAACACCTCTGGAAGCATCTCATCCAGCAT	1039
QY	1189	CCAGACAGAGTCGAGTGAAGAGATGTGCAAGATATGGAATGCCAGGAATAAAATCACT	1248
Db	1040	CCAGACAGAGTCGAGTGAAGAGATGTGCAAGATATGGAATGCCAGGAATAAAATCACT	1099
QY	1249	CAAGACATCTGCAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAGGAGTTCAGAG	1308
Db	1100	CAAGACATCTGCAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAGGAGTTCAGAG	1159
QY	1309	GACCTGCTGAGCTGCAGACGTTGACAGTGTGCTCAGACATTTAGCCAGCCATACGTC	1368
Db	1160	GACCTGCTGAGCTGCAGACGTTGACAGTGTGCTCAGACATTTAGCCAGCCATACGTC	1219
QY	1369	GTGTGCGGCGAGTCTCTGAGTACAGAGCGGCGGCGAGCCCTCCCACTGTGCCAGCA	1428
Db	1220	GTGTGCGGCGAGTCTCTGAGTACAGAGCGGCGGCGAGCCCTCCCACTGTGCCAGCA	1279
QY	1429	CCGAGGCGGAGCAGGAGCCCAAGCCCTGGGGATGCAACCTTCCACGTCCTCAGC	1488
Db	1280	CCGAGGCGGAGCAGGAGCCCAAGCCCTGGGGATGCAACCTTCCACGTCCTCAGC	1339
QY	1489	CTGACACAGCAGTCCAGGATTAAGTGTGCTTCTGCAAGGAAGCCAGCCCTGTGCACC	1548
Db	1340	CTGACACAGCAGTCCAGGATTAAGTGTGCTTCTGCAAGGAAGCCAGCCCTGTGCACC	1399
QY	1549	TGCTGCTTCAGCCATGTCGCGGAGGAGGCGGAGCGGAGCCCGGCTGTGCGC	1608
Db	1400	TGCTGCTTCAGCCATGTCGCGGAGGAGGCGGAGCGGAGCCCGGCTGTGCGC	1459
QY	1609	CCTCAGCAGTGTGCGTCTGCTGCAAGCTTTTCTGCCACCTGTACTGGGGCTGCAACCGG	1668
Db	1460	CCTCAGCAGTGTGCGTCTGCTGCAAGCTTTTCTGCCACCTGTACTGGGGCTGCAACCGG	1519
QY	1669	ACCGGTGTACGGCTGTGCGGCTGTGAGCTCAACCTGGTGCAGAGTGTCTG	1728
Db	1520	ACCGGTGTACGGCTGTGCGGCTGTGAGCTCAACCTGGTGCAGAGTGTCTG	1579
QY	1729	GACGGGTGTGAAACCAACACAGTACAGATCAGATCTGGAAGAAATACCTGGCAACC	1788

Db	1580	GACGGCGTGTGAAACAACAACAGCTACGAGTCAGACATCTCTGAAGAAATACCTGCGAACC	1639
QY	1789	AGAGGTTTTCACATGCAAAAACATGTTGACCCGAGAGCCTCTGTTGGTCTCTCCAGCGGAGTG	1848
Db	1640	AGAGGTTTTCACATGCAAAAACATGTTGACCCGAGAGCCTCTGTTGGTCTCTCCAGCGGAGTG	1699
QY	1849	TTTCTGCTGTCTGATTAACAGTCAAGGAGACACCGTCTCTGTGTTACTGCTGTGCGCTG	1908
Db	1700	TTTCTGCTGTCTGATTAACAGTCAAGGAGACACCGTCTCTGTGTTACTGCTGTGCGCTG	1759
QY	1909	CGCAGCTTCCGCTGAGCTGACCTATCAGTATCGGCAAGAAATCTCTGCTTCCGAGTTGCCA	1968
Db	1760	CGCAGCTTCCGCTGAGCTGACCTATCAGTATCGGCAAGAAATCTCTGCTTCCGAGTTGCCA	1819
QY	1969	GTGGCGGTAACTCCCGTCTCTGACTGCTACTGGGCGGTAACTGCGGCACTCAGGTGAAA	2028
Db	1820	GTGGCGGTAACTCCCGTCTCTGACTGCTACTGGGCGGTAACTGCGGCACTCAGGTGAAA	1879
QY	2029	GCTCACCAGCCCATGAATTAATCATATCTGTGAACACAGAGGTTCAAAAACCTAAGCA	2088
Db	1880	GCTCACCAGCCCATGAATTAATCATATCTGTGAACACAGAGGTTCAAAAACCTAAGCA	1939
QY	2089	TCCAGAGGCGCTGAGCAGCTTTTTCAGCATCTGAGGTTGAAGAGCGGTGTTTTTAAATACA	2148
Db	1940	TCCAGAGGCGCTGAGCAGCTTTTTCAGCATCTGAGGTTGAAGAGCGGTGTTTTTAAATACA	1999
QY	2149	GAGACAGCACCTCAAGTGTGTTTTCACAGCCCTGAGGGAAGGAGCGAGGCTCTCCCA	2208
Db	2000	GAGGCAAGCAGCTCAAGTGTGTTTTCACAGCCCTGAGGGAAGGAGCGAGGCTCTCCCA	2059
QY	2209	CAGGTGCTCTGGGGTGAAGCTTTTCTGTGAGCTTTTACCTCTGAGTGAGACCTCTCCCA	2268
Db	2060	CAGGTGCTCTGGGGTGAAGCTTTTCTGTGAGCTTTTACCTCTGAGTGAGACCTCTCCCA	2117
QY	2269	GAGCCCGGGGGCGGAGCGCCCTCTGAGGCGCTGGGCGAGGCTCTGTTGGCAT	2328
Db	2118	GAGCCCGGGGGCGGAGCGCCCTCTGAGGCGCTGGGCGAGGCTCTGTTGGCAT	2177
QY	2329	CAGCAGCAGAGACGAAAGCCTTTCTGTAAACATGCGGCGCTCCCGCCGAGAGGGGCGAGTTT	2388
Db	2178	CAGCAGCAGAGACGAAAGCCTTTCTGTAAACATGCGGCGCTCCCGCCGAGAGGGGCGAGTTT	2237
QY	2389	GCTCTTTTGTACATTTTCGAAATCTAGTTAAAGCAGAGTCTGTTTTCAGGAAAGTT	2448
Db	2238	GCTCTTTTGTACATTTTCGAAATCTAGTTAAAGCAGAGTCTGTTTTCAGGAAAGTT	2297
QY	2449	TCAAGGAGAGGCGCAAGTTTATCAAAAACATTTGTTTTCAGGAGAGGAGGAGCATAGTTTA	2508
Db	2298	TCAAGGAGAGGCGCAAGTTTATCAAAAACATTTGTTTTCAGGAGAGGAGGAGCATAGTTTA	2357
QY	2509	CAGCTACAGGACGTACAAATATCTGCTGCTGGGAAACACACAGCATTCTATCTATTT	2568
Db	2358	CAGCTACAGGACGTACAAATATCTGCTGCTGGGAAACACACAGCATTCTATCTATTT	2417
QY	2569	TTTATTTTAAATAGGTTTGGTGTCTTCTCTAATAAGATTAAATGTCAAAAACCTGTAGC	2628
Db	2418	TTTATTTTAAATAGGTTTGGTGTCTTCTCTAATAAGATTAAATGTCAAAAACCTGTAGC	2477
QY	2629	ACAAATTAATAATTTAATAATTTCAAAATGAC 2661	
Db	2478	ACAAATTAATAATTTAATAATTTCAAAATGAC 2510	

RESULT 6
AAI88903/c
ID AAI88903 standard; cDNA; 2259 BP.
XX
AC AAI88903;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 8963.
XX

XX Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
XX preservative; antiproliferative.
OS Homo sapiens.
XX
XX WO20015322-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216547P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228242P.
PR 01-SEP-2000; 2000US-0229487P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 23-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.

14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
22-AUG-2000; 2000US-0228686P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229503P.
05-SEP-2000; 2000US-0229513P.
08-SEP-2000; 2000US-0231413P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
27-SEP-2000; 2000US-0235834P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239335P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241803P.
01-NOV-2000; 2000US-0244617P.
17-NOV-2000; 2000US-0249299P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;
WPI; 2003-147444/14.
P-PSDB; ABU54924.

New polypeptides and nucleic acids, useful in gene therapy for treating, inhibiting or preventing e.g. neural, immune system, muscular, respiratory, reproductive, Gastrointestinal, pulmonary, cardiovascular or renal disorders.

Claim 1; SEQ ID NO 22; 402pp; English.

The invention relates to human novel polypeptides and their associated polynucleotides. The polypeptides and polynucleotides are useful in gene therapy for treating, inhibiting or preventing neural disorders, immune system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis), muscular disorders, respiratory diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders, gastrointestinal disorders, pulmonary disorders, cardiovascular disorders, (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute kidney failure and end-stage renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and leukaemia), inflammatory diseases (e.g. septic shock, bursitis and appendicitis), allergic reactions and conditions (e.g. asthma), blood related disorders (e.g. thrombosis, atherosclerosis and myocardial infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent human novel polynucleotides of the invention

Query Match 28.6%; Score 765; DB 7; Length 1311;
Best Local Similarity 100.0%; Pred. No. 2.2e-273;
Matches 765; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 710 GGTCTGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAAGTGAAG 769
DB 383 GGTCTGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAAGTGAAG 442
QY 770 TCTCCAGCTTTGCTCAGCTCTCCAGACAGAAAGACTCGCTCTTTCGTGTTGGAAC 829
DB 443 TCTCCAGCTTTGCTCAGCTCTCCAGACAGAAAGACTCGCTCTTTCGTGTTGGAAC 502
QY 830 CCCAGGATCAGGAGGATTTGGAGCCGCTGAAGAGAAATGAGAGAGATGGGACCTTG 889
DB 503 CCCAGGATCAGGAGGATTTGGAGCCGCTGAAGAGAAATGAGAGAGATGGGACCTTG 562
QY 890 ACCTGAACCGGCGAGTTTGGTGGCAAAACCGCGTAGAAATGCCAAAACCGTCCAGAG 949
DB 563 ACCTGAACCGGCGAGTTTGGTGGCAAAACCGCGTAGAAATGCCAAAACCGTCCAGAG 622
QY 950 ACCTGAACCGGCGAGTTTGGTGGCAAAACCGCGTAGAAATGCCAAAACCGTCCAGAG 1009
DB 623 ACCTGAACCGGCGAGTTTGGTGGCAAAACCGCGTAGAAATGCCAAAACCGTCCAGAG 582
QY 1010 GCCAGGACCTGTGACGAGTTCGCTGAGTTTGCAGCCCTGCATGCACACGTTCTGGCGG 1069
DB 683 GCCAGGACCTGTGACGAGTTCGCTGAGTTTGCAGCCCTGCATGCACACGTTCTGGCGG 742
QY 1070 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCTTACCTGCGCGTGTCCCGTGG 1129
DB 743 CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCTTACCTGCGCGTGTCCCGTGG 802
QY 1130 AGCGGATCTGTAAAAACCAACATCTCTCAACAACTCGTGAAGCATACCTCATCCAGCATC 1189
DB 803 AGCGGATCTGTAAAAACCAACATCTCTCAACAACTCGTGAAGCATACCTCATCCAGCATC 862
QY 1190 CAGACAGAGTGCAGTGAAGAGATGTGCAAGATAGTATGATGCCAGGAATAAATCAGTTC 1249
DB 863 CAGACAGAGTGCAGTGAAGAGATGTGCAAGATAGTATGATGCCAGGAATAAATCAGTTC 922
QY 1250 AAGACATGTCTGAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAGGAGTTCAGAGG 1309
DB 923 AAGACATGTCTGAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAGGAGTTCAGAGG 982
QY 1310 ACCTGTGAGAGTGTGACAGCTTTCAGAGTTCCTCAGACATTAGCCAGCCATAGTTCG 1369
DB 983 ACCTGTGAGAGTGTGACAGCTTTCAGAGTTCCTCAGACATTAGCCAGCCATAGTTCG 1042
QY 1370 TGTGCGGCGAGTGTCTCTGAGTACAGAAAGCAGCGCGCAGAGCTCCCTCCCTAGCCAGCAC 1429
DB 1043 TGTGCGGCGAGTGTCTCTGAGTACAGAAAGCAGCGCGCAGAGCTCCCTCCCTAGCCAGCAC 1102
QY 1430 CCGAGGCGGAGCCAGAGCCCAACAGGCGCTGGGGGATGCACCCCT 1474
DB 1103 CCGAGGCGGAGCCAGAGCCCAACAGGCGCTGGGGGATGCACCCCT 1147

RESULT 9
ADA52592
ID ADA52592 standard; cDNA; 2186 BP.
XX
AC ADA52592;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human coding sequence, SEQ ID 160.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.
OS Homo sapiens.
XX
FN EP1293569-A2.
XX
PD 19-MAR-2003.

Db 361 TGGGGATGTCTACTTGGTGTACAGGAAGTAACCGGAACACACAGTGCATACCT 420
QY 447 CTATGAATCTTTAAGTGAAGCAAGGCGATGACACAAAGTCTTTGA 494
Db 421 CTATGAATCTTTAAGTGAAGCAAGGCGATGACACAAAGTCTTTGA 469

RESULT 12
AAS26304
ID AAS26304 standard; cDNA; 693 BP.
AC AAS26304;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 483.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nontropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216847P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 28-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232050P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233083P.
PR 14-SEP-2000; 2000US-0233084P.
PR 14-SEP-2000; 2000US-0233085P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249220P.
PR 17-NOV-2000; 2000US-0249221P.
PR 17-NOV-2000; 2000US-0249222P.
PR 17-NOV-2000; 2000US-0249223P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249225P.
PR 17-NOV-2000; 2000US-0249226P.
PR 17-NOV-2000; 2000US-0249227P.
PR 17-NOV-2000; 2000US-0249228P.
PR 17-NOV-2000; 2000US-0249229P.
PR 17-NOV-2000; 2000US-0249230P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0253678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-488783/53.
XX P-PSDB; AAU16317.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
XX diagnosing, preventing, treating or ameliorating medical conditions and
XX used as food additives or preservatives.
XX
XX Claim 1; SEQ ID NO 483; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
XX encoded secreted proteins. The nucleic acids and proteins are used to
XX prevent, treat or ameliorate a medical condition in e.g. humans, mice,
XX rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
XX in diagnosing a pathological condition or susceptibility to a
XX pathological condition. Antibodies to the proteins can also be used in
XX alleviating symptoms associated with the disorders and in diagnostic
XX immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
XX (ELISA). Disorders which are diagnosed or treated include autoimmune
XX diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
XX neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
XX arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
XX nervous system disorders e.g. Alzheimer's disease, infections caused by
XX bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
XX and many other disorders listed in the specification. The polypeptides
XX can also be used to aid wound healing and epithelial cell proliferation,
XX to prevent skin aging due to sunburn, to maintain organs before
XX transplantation, for supporting cell culture of primary tissues, to
XX regenerate tissues and in chemotaxis. The polypeptides can also be used
XX as a food additive or preservative to increase or decrease storage
XX capabilities, fat content, lipid, protein, carbohydrate, vitamins,
XX minerals, cofactors and other nutritional components. The present
XX sequence encodes a novel secreted protein of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX
XX Query Match 17.0%; Score 455; DB 4; Length 693;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-158; Gaps 0;
XX Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1044 GCCTGATGACACAGCTTCTGCGGCTTCTACTCTGCGGCTGATGAGGCTGCTCCCT 1103
XX
XX 14 GCCTGATGACACAGCTTCTGCGGCTTCTACTCTGCGGCTGATGAGGCTGCTCCCT 73
XX
XX 1104 GTGTCTACCTGCGGCTGTCCTGCGGAGGATCTGTAAAAACCATCTCTCAACACCT 1163

Db 74 GTGTCTACCTGCGGCTGTCCTGCGGAGCGGATCTGTAAAAACCATCTCTCAACACCT 133
Qy 1164 CGTGAAGCATACCTCATCCAGCATCCAGACAAGAGTCCAGTGAAGAAGATGTGCAAG 1223
Db 134 CGTGAAGCATACCTCATCCAGCATCCAGACAAGAGTCCAGTGAAGAAGATGTGCAAG 193
Qy 1224 TATGATGCCAGGAATAAATCACTCAAGCATCTGTCGAGCCCAAGAGTCAGGCGTCTTT 1283
Db 194 TATGATGCCAGGAATAAATCACTCAAGCATCTGTCGAGCCCAAGAGTCAGGCGTCTTT 253
Qy 1284 TTCTGATGAAGAGGAGTTCAGAGGACCTGTCGAGCTGTGAGAGCTTTCAGAGTCAGTC 1343
Db 254 TTCTGATGAAGAGGAGTTCAGAGGACCTGTCGAGCTGTGAGAGCTTTCAGAGTCAGTC 313
Qy 1344 CTCAGACATTAGCCAGCCATACGTCTGTGCCGCGAGTGTCTTGTAGTACAGAAGGAGGC 1403
Db 314 CTCAGACATTAGCCAGCCATACGTCTGTGCCGCGAGTGTCTTGTAGTACAGAAGGAGGC 373
Qy 1404 GCGCGAGCTCCCTCCACTGTCGAGCACCAGGAGGCGAGCCAGAGCCCAAGAGCTTGGG 1463
Db 374 GCGCGAGCTCCCTCCACTGTCGAGCACCAGGAGGCGAGCCAGAGCCCAAGAGCTTGGG 433
Qy 1464 GGATGCACCTCCCTCCACTGTCGAGCACCAGGAGGCGAGCCAGAGCTTGGGACAG 1498
Db 434 GGATGCACCTCCCTCCACTGTCGAGCACCAGGAGGCGAGCTTGGGACAG 468
RESULT 13
ABX73645
ID ABX73645 standard; DNA; 693 BP.
XX AC ABX73645;
XX
XX 18-MAR-2003 (first entry)
XX
XX Human novel polynucleotide #473.
XX
XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;
XX muscular disorder; respiratory disease; reproductive disorder;
XX gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX hyperproliferative disorder; inflammatory disease; allergic reaction;
XX blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
XX haemostatic; antiarteriosclerotic.
XX
XX Homo sapiens.
XX
XX US2002132753-A1.
XX
XX 19-SEP-2002.
XX
XX 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 28-JUN-2000; 2000US-0214886P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220984P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 22-AUG-2000; 2000US-0226868P.

PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
PI
XX WPI; 2003-147444/14.
DR P-PSDB; ABUS5385.
DR
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 483; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention.
XX
SQ Sequence 693 BP; 145 A; 214 C; 197 G; 123 T; 0 U; 14 Other;
Query Match 17.0%; Score 455; DB 7; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.3e-158;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1044 GCCTGATGACACAGTCTCGCGGCTTCTACTCGCGGTGATGAGCGGCTCGCCCT 1103
Db 14 GCCTGATGACACAGTCTCGCGGCTTCTACTCGCGGTGATGAGCGGCTCGCCCT 73

QY 1104 GTGTCCTACTGCGCGCTGTCCCGTGGAGCGGATCTGTAAAAACACATCTCTCAACACCT 1163
Db 74 GTGTCCTACTGCGCGCTGTCCCGTGGAGCGGATCTGTAAAAACACATCTCTCAACACCT 133
QY 1164 CGTGAAGCATACCTCATCTCCAGCATCCAGACAAGATCGCAGTGAAGAAGATGTGCAAG 1223
Db 134 CGTGAAGCATACCTCATCTCCAGCATCCAGACAAGATCGCAGTGAAGAAGATGTGCAAG 193
QY 1224 TATGATGCCAGGAATAAATCACTCAAGACATCTGAGGCCCAAGTCAGGCGGTCTTT 1283
Db 194 TATGATGCCAGGAATAAATCACTCAAGACATCTGAGGCCCAAGTCAGGCGGTCTTT 253
QY 1284 TTCTGATGAAGAGGAGTTTCAGAGGACCTCTCGAGTGTGACAGCTTGACAGTGAGTC 1343
Db 254 TTCTGATGAAGAGGAGTTTCAGAGGACCTCTCGAGTGTGACAGCTTGACAGTGAGTC 313
QY 1344 CTCAGACATTAGCCAGGCCATACGTGTCGCGGAGTGTCTGAGTACAGAAAGCAGGC 1403
Db 314 CTCAGACATTAGCCAGGCCATACGTGTCGCGGAGTGTCTGAGTACAGAAAGCAGGC 373
QY 1404 GCGCAGGCTCCCGACCTGCCAGCACCCGAGGCGAGCCAGAGGCCCCACAGGCCCTGGG 1463
Db 374 GCGCAGGCTCCCGACCTGCCAGCACCCGAGGCGAGCCAGAGGCCCCACAGGCCCTGGG 433
QY 1464 GGATGCACCCCTCCAGCTCCGTCAGCTGACGACAG 1498
Db 434 GGATGCACCCCTCCAGCTCCGTCAGCTGACGACAG 468

RESULT 14
AAH11859/c
ID AAH11859 standard; cDNA; 518 BP.
XX AC AAH11859;
XX DT 26-JUN-2001 (first entry)
XX DE Human cDNA clone (3'-primer) SEQ ID NO:8694.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX FN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 93JP-00248036.
XX PR 27-AUG-1999; 93JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX Claim 3; SEQ ID NO 8694; 2537pp + Sequence Listing; English.
XX CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the

complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 518 BP; 142 A; 127 C; 108 G; 136 T; 0 U; 5 Other;

Query Match 15.9%; Score 426; DB 4; Length 518;
Best Local Similarity 99.8%; Pred. No. 7.2e-148;
Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2185 AGGGAAGGACGACGGGTCTCCGACAGGTGCTCTGGGTGACTCTCTGTGAGCTTTT 2244
DB 503 AGGGAAGGACGACGGGTCTCCGACAGGTGCTCTGGGTGACTCTCTGTGAGCTTTT 444
QY 2245 ACCCTCTGAGTGAGACCTCCGACAGCCCGGGGGCCGACGCCCTCTGTGAGC 2304
DB 443 ACCCTCTGAGTGAGACCTCCGACAGCCCGGGGGCCGACGCCCTCTGTGAGC 384
QY 2305 GCTGGGACGGGCTCTGGGTGGCATCAGCAGCAGAGCAAGCTTTCTGTAACTGCGC 2364
DB 383 GCTGGGACGGGCTCTGGGTGGCATCAGCAGCAGAGCAAGCTTTCTGTAACTGCGC 324
QY 2365 CGTCCCGCCGAGAGGGGAGTTTGTCTTTGTGTATCTTCGAACTACAGTTAAGC 2424
DB 323 CGTCCCGCCGAGAGGGGAGTTTGTCTTTGTGTATCTTCGAACTACAGTTAAGC 264
QY 2425 AGAAGTCTGTTTTCAGGAAAGTTTCAAGGAGAGGCAAGTTATCAAAACATTTGT 2484
DB 263 AAAAGTCTGTTTTCAGGAAAGTTTCAAGGAGAGGCAAGTTATCAAAACATTTGT 204
QY 2485 TCAGGAGAGGGAGCAATAGTTTACAGCTTACAGGACGTACCAATATCTCTGCTGGG 2544
DB 203 TCAGGAGAGGGAGCAATAGTTTACAGCTTACAGGACGTACCAATATCTCTGCTGGG 144
QY 2545 AAAACCCACAGCATTTATCTATTTTATTTTATAGTTTGTGTTTCTTCTTAATAA 2604
DB 143 AAAACCCACAGCATTTATCTATTTTATTTTATAGTTTGTGTTTCTTCTTAATAA 84
QY 2605 GATTTAATGTCACAACTGTAGCACAATAATAATAATAATAATAATAATAATAATA 2661
DB 83 GATTTAATGTCACAACTGTAGCACAATAATAATAATAATAATAATAATAATAATA 27

RESULT 15
AAA44336
ID AAA44336 standard; cDNA; 575 BP.

XX AC AAA44336;

XX DT 21-AUG-2000 (first entry)

XX Human secreted expressed sequence tag SEQ ID NO:911.

XX Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
XX expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;

antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;
antulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;
autoimmune disorder; multiple sclerosis; allergic condition;
insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
central nervous system disorder; Alzheimer's disease; stroke;
Parkinson's disease; Huntington's disease; coagulation disorder;
haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;
infection; depression; psoriasis; ss.

XX Homo sapiens.

OS WO200021991-A1.

PN 20-APR-2000.

PD 15-OCT-1999; 99WO-US024206.

PR 15-OCT-1998; 98US-0104436P.

XX (GEMY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Werberg D, Treacy M, Bowman MR;

XX WPI; 2000-317938/27.

Isolated polynucleotides, and encoded proteins, comprising secreted expressed sequence tags (sESTs), useful for treating various disorders such as autoimmune, infectious, and central nervous system disorders.

Claim 1; Page 437; 803pp; English.

AAA43426 to AAA45925 represent specifically claimed secreted expressed sequence tags (sESTs), isolated from human, mouse, chicken and rat tissue sources. The sESTs can have a range of activities depending on the tissues they were isolated from. The activities include: chemotactic; proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal; antiviral; antidiabetic; antiasthmatic; vulnary; antulcer; osteopathic; neuroprotective; nootropic; antiparkinsonian; antipsoriatic; cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be used for gene therapy and in vaccines. The sESTs are useful as probes for the identification and isolation of full-length cDNAs and genomic DNA molecules which correspond to the sESTs. Proteins encoded by the sESTs are useful in assays for determining biological activity and raising antibodies. They may be useful for treatment of autoimmune disorders (multiple sclerosis, insulin dependent diabetes), allergic conditions (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers, osteoporosis, osteoarthritis, central nervous system disorders (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's disease), tumours, bacterial fungal or viral infections, depression and psoriasis. AAA45926 to AAA45931 represent linker variants which are given in the exemplification of the present invention

XX Sequence 575 BP; 124 A; 168 C; 165 G; 118 T; 0 U; 0 Other;

Query Match 10.7%; Score 286; DB 3; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.5e-96;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 876 AGATGGGACCTTGACCTGAACGGGCGAGTTGTGGTCGACACACCGCGTAGAATGCCA 935

DB 285 AGATGGGACCTTGACCTGAACGGGCGAGTTGTGGTCGACACACCGCGTAGAATGCCA 344

QY 936 AACGTCACAGGACGTCAGAGCAGCGGCTGGGAAGCCAGACAGATGGAGGAGCGCT 995

DB 345 AACGTCACAGGACGTCAGAGCAGCGGCTGGGAAGCCAGACAGATGGAGGAGCGCT 404

QY 996 GACATGCATCTCTGCCAGGACCTGCTGCACGATGCTGTGAGTTTGAGCCCTCATGCA 1055

ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: hazel (Corylus sp.)
IMMEDIATE SOURCE:
LIBRARY: POLLEN FROM ALLERSON AB, ENGELHOLM, SWEDEN
US-07-847-010-18

Query Match
Best Local Similarity 1.0%; Score 26; DB 1; Length 860;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2654 AATTGACAAAAA 2679
DB 754 AATTGACAAAAA 729

RESULT 2
US-09-522-714-15
Sequence 15, Application US/09522714
Patent No. 6563020
GENERAL INFORMATION:
APPLICANT: Simmons, Carl R.
APPLICANT: Valpani, Nasser
TITLE OF INVENTION: Maize Chitinases and Their Use in
TITLE OF INVENTION: Enhancing Disease Resistance in Crop Plants
FILE REFERENCE: 1100
CURRENT APPLICATION NUMBER: US/09/522,714
CURRENT FILING DATE: 2000-03-10
EARLIER APPLICATION NUMBER: 60/125,915
EARLIER FILING DATE: 1999-03-24
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 1085
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (46)....(966)
US-09-522-714-15

Query Match
Best Local Similarity 0.9%; Score 25; DB 4; Length 1085;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679
DB 1061 AATTGACAAAAA 1085

RESULT 3
5198347-3
Patent No. 5198347
APPLICANT: Miller, Louis H.; Adams, John H.; Kaslow,
DAVID C.; Fang, Xianguo
TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
PLASMODIUM KNOWLES DUFFY RECEPTOR
NUMBER OF SEQUENCES: 27
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,837
FILING DATE: 20-JUL-1990
SEQ ID NO: 3
LENGTH: 3157
5198347-3

Query Match
Best Local Similarity 0.9%; Score 24; DB 6; Length 3157;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679
DB 3088 ATTGACAAAAA 3111

RESULT 4
US-09-621-976-17572
Sequence 17572, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 17572
LENGTH: 443
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-17572

Query Match
Best Local Similarity 0.9%; Score 23; DB 4; Length 443;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
DB 420 TTGACAAAAA 442

RESULT 5
US-09-614-912-161
Sequence 161, Application US/09614912
Patent No. 6677502
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalski, Antoni
APPLICANT: Orozco, Buddy
APPLICANT: Miao, Gou-Hau
APPLICANT: Pamodu, Omolayo O.
APPLICANT: Lee, Jian Ming
APPLICANT: Sakai, Hajime
APPLICANT: Meng, Zude
APPLICANT: Cai, Perry G
APPLICANT: Anderson, Shawn
TITLE OF INVENTION: Plant Metabolism Genes
FILE REFERENCE: BB1378 US NA
CURRENT APPLICATION NUMBER: US/09/614,912
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 60/143,401
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/143,412
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/146,650
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/170,906
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: 60/172,959
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/172,946
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 204
SOFTWARE: Microsoft Office 97
SEQ ID NO 161
LENGTH: 1185
TYPE: DNA
ORGANISM: Glycine max
US-09-614-912-161

Query Match
Best Local Similarity 0.9%; Score 23; DB 4; Length 1185;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: hazel (Corylus sp.)
IMMEDIATE SOURCE:
LIBRARY: POLLEN FROM ALLERSON AB, ENGELHOLM, SWEDEN
US-07-847-010-18

Query Match
Best Local Similarity 1.0%; Score 26; DB 1; Length 860;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2654 AATTGACAAAAA 2679
DB 754 AATTGACAAAAA 729

RESULT 2
US-09-522-714-15
Sequence 15, Application US/09522714
Patent No. 6563020
GENERAL INFORMATION:
APPLICANT: Simmons, Carl R.
APPLICANT: Valpani, Nasser
TITLE OF INVENTION: Maize Chitinases and Their Use in
TITLE OF INVENTION: Enhancing Disease Resistance in Crop Plants
FILE REFERENCE: 1100
CURRENT APPLICATION NUMBER: US/09/522,714
CURRENT FILING DATE: 2000-03-10
EARLIER APPLICATION NUMBER: 60/125,915
EARLIER FILING DATE: 1999-03-24
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 1085
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (46)....(966)
US-09-522-714-15

Query Match
Best Local Similarity 0.9%; Score 25; DB 4; Length 1085;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679
DB 1061 AATTGACAAAAA 1085

RESULT 3
5198347-3
Patent No. 5198347
APPLICANT: Miller, Louis H.; Adams, John H.; Kaslow,
DAVID C.; Fang, Xianguo
TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
PLASMODIUM KNOWLES DUFFY RECEPTOR
NUMBER OF SEQUENCES: 27
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,837
FILING DATE: 20-JUL-1990
SEQ ID NO: 3
LENGTH: 3157
5198347-3

Query Match
Best Local Similarity 0.9%; Score 24; DB 6; Length 3157;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679
DB 3088 ATTGACAAAAA 3111

RESULT 4
US-09-621-976-17572
Sequence 17572, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 17572
LENGTH: 443
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-17572

Query Match
Best Local Similarity 0.9%; Score 23; DB 4; Length 443;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
DB 420 TTGACAAAAA 442

RESULT 5
US-09-614-912-161
Sequence 161, Application US/09614912
Patent No. 6677502
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalski, Antoni
APPLICANT: Orozco, Buddy
APPLICANT: Miao, Gou-Hau
APPLICANT: Pamodu, Omolayo O.
APPLICANT: Lee, Jian Ming
APPLICANT: Sakai, Hajime
APPLICANT: Meng, Zude
APPLICANT: Cai, Perry G
APPLICANT: Anderson, Shawn
TITLE OF INVENTION: Plant Metabolism Genes
FILE REFERENCE: BB1378 US NA
CURRENT APPLICATION NUMBER: US/09/614,912
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 60/143,401
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/143,412
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/146,650
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/170,906
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: 60/172,959
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/172,946
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 204
SOFTWARE: Microsoft Office 97
SEQ ID NO 161
LENGTH: 1185
TYPE: DNA
ORGANISM: Glycine max
US-09-614-912-161

Query Match
Best Local Similarity 0.9%; Score 23; DB 4; Length 1185;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
Db 1163 TTGACAAAAA 1185

RESULT 6
US-09-412-102-1
; Sequence 1, Application US/09412102
; Patent No. 6228992
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY J
; APPLICANT: MEYER, TERRY E
; TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF
; NEMATODES IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: US
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412,102
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/217,787
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-18P1
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1347 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; IMMEDIATE SOURCE:
; CLONE: P12217
; NAME/KEY: CDS
; LOCATION: 146..994
; US-09-412-102-1

Query Match 0.9%; Score 23; DB 3; Length 1347;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
Db 1325 TTGACAAAAA 1347

RESULT 7
US-09-217-787-1
; Sequence 1, Application US/09217787
; Patent No. 6284948
; GENERAL INFORMATION:
; APPLICANT: JESSEN, HOLLY J
; APPLICANT: MEYER, TERRY E

; TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF
; NEMATODES IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: US
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/217,787
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-18P1
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1347 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; IMMEDIATE SOURCE:
; CLONE: P12217
; NAME/KEY: CDS
; LOCATION: 146..994
; US-09-217-787-1

Query Match 0.9%; Score 23; DB 3; Length 1347;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
Db 1325 TTGACAAAAA 1347

RESULT 8
US-08-484-101B-35
; Sequence 35, Application US/08484101B
; Patent No. 5824868
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO
; ETHYLENE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:


```
/ APPLICATION NUMBER: US/08/484,101B
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA: PCT/US94/
/ FILING DATE: 01-JUL-1994
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/086,555
/ FILING DATE: 01-JUL-1993
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Trecartin, Richard F.
/ REGISTRATION NUMBER: 31,801
/ REFERENCE/DOCKET NUMBER: A-57515-2/RFT
/ TELEPHONE: (415) 781-1989
/ TELEFAX: (415) 398-3249
/ INFORMATION FOR SEQ ID NO: 35:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2405 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 288..2196
/ US-08-484-101B-35

Query Match 0.9%; Score 23; DB 1; Length 2405;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
Db 2378 TTGACAAAAA 2400

RESULT 9
US-08-484-101B-49
/ Sequence 49, Application US/08/484101B
/ Patent No. 5824868
/ GENERAL INFORMATION:
/ APPLICANT: California Institute of Technology
/ TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO
/ TITLE OF INVENTION: ETHYLENE
/ NUMBER OF SEQUENCES: 50
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Richard F. Trecartin
/ STREET: 3400 Embarcadero Center, Suite 3400
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC Compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/484,101B
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US94/
/ FILING DATE: 01-JUL-1994
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/086,555
/ FILING DATE: 01-JUL-1993
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Trecartin, Richard F.

/ APPLICATION NUMBER: 31,801
/ REFERENCE/DOCKET NUMBER: A-57515-2/RFT
/ TELEPHONE: (415) 781-1989
/ TELEFAX: (415) 398-3249
/ INFORMATION FOR SEQ ID NO: 49:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2405 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 288..2196
/ US-08-484-101B-49

Query Match 0.9%; Score 23; DB 1; Length 2405;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
Db 2378 TTGACAAAAA 2400

RESULT 10
US-08-714-524D-35
/ Sequence 35, Application US/08/714524D
/ Patent No. 6294716
/ GENERAL INFORMATION:
/ APPLICANT: Meyerowitz, Elliott M
/ APPLICANT: Chang, Caren
/ APPLICANT: Bleeker, Anthony B
/ TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
/ FILE REFERENCE: a-57515-4
/ CURRENT APPLICATION NUMBER: US/08/714,524D
/ CURRENT FILING DATE: 1996-09-16
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 35
/ LENGTH: 2405
/ TYPE: DNA
/ ORGANISM: Lycopersicon esculentum
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (288)..(2195)
/ US-08-714-524D-35

Query Match 0.9%; Score 23; DB 3; Length 2405;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
Db 2378 TTGACAAAAA 2400

RESULT 11
US-08-714-524D-49
/ Sequence 49, Application US/08/714524D
/ Patent No. 6294716
/ GENERAL INFORMATION:
/ APPLICANT: Meyerowitz, Elliott M
/ APPLICANT: Chang, Caren
/ APPLICANT: Bleeker, Anthony B
/ TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
/ FILE REFERENCE: a-57515-4
/ CURRENT APPLICATION NUMBER: US/08/714,524D
/ CURRENT FILING DATE: 1996-09-16
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 49
/ LENGTH: 2405
```

; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (288)..(2195)
US-08-714-524D-49

Query Match 0.9%; Score 23; DB 3; Length 2405;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679
DB 2378 TTGACAAAAA 2400

RESULT 12

US-08-244-113-18
; Sequence 18, Application US/08244113
; Patent No. 5455181

GENERAL INFORMATION:

; APPLICANT: Strube, Karl-Hermann
; APPLICANT: Bialojan, Siegfried
; APPLICANT: Kroege, Burkhard
; APPLICANT: Friedrich, Thomas
; TITLE OF INVENTION: No. 5455181el thrombin-inhibitory proteins from terrestrial
; TITLE OF INVENTION: leeches.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Keil & Weinkauff
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.0
; SOFTWARE: WordPerfect version 5.1

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/244,113
; FILING DATE:

CLASSIFICATION:

; CLASSIFICATION: C07K 73/10
; CLASSIFICATION: A61K 37/64

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP92/02661

FILING DATE:

; FILING DATE: 19-NOV-1992

INFORMATION FOR SEQ ID NO:

; INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

; LENGTH: 277 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-244-113-18

Query Match 0.8%; Score 22; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAA 2679
DB 224 TGACAAAAA 245

RESULT 13

US-09-280-116-159/c
; Sequence 159, Application US/09280116A
; Patent No. 6331427

GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs

; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 159
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Lon family of ATP-dependent proteases
US-09-280-116-159

Query Match 0.8%; Score 22; DB 4; Length 424;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAA 2679
DB 22 TGACAAAAA 1

RESULT 14

US-09-205-258-172

; Sequence 172, Application US/09205258

; Patent No. 6525174

GENERAL INFORMATION:

; APPLICANT: Young et al.

; TITLE OF INVENTION: 207 Human Secreted Proteins

; FILE REFERENCE: P2007P1

; CURRENT APPLICATION NUMBER: US/09/205,258

; CURRENT FILING DATE: 1998-12-04

; EARLIER APPLICATION NUMBER: PCT/US98/11422

; EARLIER FILING DATE: 1998-06-04

; EARLIER APPLICATION NUMBER: 60/048,885

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/049,375

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,881

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,880

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,896

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/049,020

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,876

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,895

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,884

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,894

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,971

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,964

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,882

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,899

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,893

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,900

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,901

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,892

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,915

; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/049,019

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,891

Query Match 0.8%; Score 22; DB 4; Length 786;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TCACAAAAA 2679
DB 760 TCACAAAAA 781

RESULT 15
US-09-205-258-190/c
Sequence 190; Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,891

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SEQ ID NO 172
LENGTH: 786
TYPE: DNA
ORGANISM: Homo sapiens
US-09-205-258-172

```

; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 190
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (144)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-190

Query Match      0.8%; Score 22; DB 4; Length 906;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2657 TTGACAAAAA 2678
      |||||
Db      843 TTGACAAAAA 822

Search completed: May 15, 2004, 08:02:31
Job time : 216 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:49:13 ; Search time 11.5628 Seconds
(without alignments)
1752.371 Million cell updates/sec

Title: US-10-048-046-2_COPY_31_103

Perfect score: 378
Sequence: 1 VLLRKEWTIGRRGCDLSF.....INKLKVVKQTCPLQGTGDI 73

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	378	100.0	426	9	US-09-764-864-809
2	378	100.0	664	9	US-09-780-525-2
3	304	80.4	92	9	US-09-764-864-1109
4	275	72.8	99	9	US-09-764-864-1110
5	275	72.8	99	9	US-09-764-864-1529
6	107.5	28.4	522	9	US-09-740-627-11
7	107.5	28.4	821	14	US-10-081-119-16
8	99	26.2	699	14	US-10-032-585-7635
9	81	21.4	445	12	US-10-424-599-178873
10	81	21.4	1082	12	US-10-425-114-56100
11	78.5	20.8	175	9	US-09-867-550-1026
12	78.5	20.8	244	9	US-09-764-864-1111
13	77.5	20.2	460	9	US-09-740-627-10
14	76.5	20.2	517	15	US-10-618-173-2
15	74.5	19.7	229	12	US-10-424-599-187236

16	74.5	19.7	376	12	US-10-424-599-233512	Sequence 233512, App
17	74	19.6	250	12	US-10-424-599-269525	Sequence 269525, App
18	72	19.0	513	15	US-10-369-493-1578	Sequence 1578, App
19	71	18.8	409	12	US-10-282-122A-64192	Sequence 64192, A
20	70.5	18.7	874	12	US-10-282-122A-45185	Sequence 45185, A
21	67	17.7	481	14	US-10-156-761-12906	Sequence 12906, A
22	66.5	17.6	288	9	US-09-738-626-3552	Sequence 3552, App
23	66	17.5	435	12	US-10-424-599-276420	Sequence 276420, App
24	66	16.9	545	14	US-10-142-356-6	Sequence 6, Appli
25	63.5	16.8	474	15	US-10-369-493-2299	Sequence 2299, App
26	63.5	16.8	492	15	US-10-369-493-3607	Sequence 3607, App
27	63	16.7	138	14	US-10-238-075-1035	Sequence 1035, App
28	62.5	16.5	527	9	US-09-712-363-156	Sequence 156, App
29	62.5	16.5	527	14	US-10-080-170-348	Sequence 348, App
30	62	16.4	467	12	US-10-406-031-5	Sequence 5, Appli
31	62	16.4	1724	12	US-10-276-774-2176	Sequence 2176, App
32	61.5	16.3	207	12	US-10-282-122A-49959	Sequence 49959, A
33	61.5	16.3	459	9	US-09-740-627-9	Sequence 9, Appli
34	61.5	16.3	488	14	US-10-050-170-4	Sequence 4, Appli
35	61	16.1	1055	12	US-10-282-122A-57704	Sequence 57704, A
36	61	16.1	1056	12	US-10-282-122A-57499	Sequence 57499, A
37	60.5	16.0	365	9	US-09-738-636-4872	Sequence 4872, App
38	60.5	16.0	397	12	US-10-389-647-430	Sequence 430, App
39	60	15.9	939	12	US-10-282-122A-47696	Sequence 47696, A
40	60	15.9	969	12	US-10-282-122A-63852	Sequence 63852, A
41	59.5	15.7	804	12	US-10-282-122A-42499	Sequence 42499, A
42	59.5	15.7	820	9	US-09-815-242-10771	Sequence 10771, A
43	59	15.6	601	12	US-10-257-502-367	Sequence 367, App
44	59	15.6	1045	12	US-10-041-018-199	Sequence 199, App
45	59	15.6	1045	12	US-10-041-018-218	Sequence 218, App

ALIGNMENTS

RESULT 1
US-09-764-864-809
Sequence 809, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 809

LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (414)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (415)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (420)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-809

Query Match 100.0%; Score 378; DB 9; Length 426;
Best Local Similarity 100.0%; Pred. No. 6.9e-39;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLRKEWTIGRRGCDLSFPSNKLVSGDHCRIVVDEKSGQVTLSDTSGTVINKLV 60
DB 33 VLLRKEWTIGRRGCDLSFPSNKLVSGDHCRIVVDEKSGQVTLSDTSGTVINKLV 92
QY 61 KKQTCPLQGTGDI 73

not 102?
date of 2000
601

```
Db      93 KKQTCPLQTGDVI 105
|||||
RESULT 2
US-09-780-525-2
; Sequence 2, Application US/09780525
; Patent No. US2002004223A1
; GENERAL INFORMATION:
; APPLICANT: Bin-Bing Zhou
; APPLICANT: Yuan Zhu
; APPLICANT: Priya Chaturvedi
; APPLICANT: Mark R. Hurler
; APPLICANT: Xiaotong Li
; TITLE OF INVENTION: FHAR1, A NEW RING FINGER PROTEIN
; FILE REFERENCE: GP-70668-C1
; CURRENT APPLICATION NUMBER: US/09/780,525
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 09/456,876
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 664
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-780-525-2
Query Match      100.0%; Score 378; DB 9; Length 664;
Best Local Similarity 100.0%; Pred. No. 1.2e-38;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 VLLRKEWTIGRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTDTSGTVINKLVV 60
Db      31 VLLRKEWTIGRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTDTSGTVINKLVV 90
Qy      61 KKQTCPLQTGDVI 73
Db      91 KKQTCPLQTGDVI 103
|||||
RESULT 3
US-09-764-864-1109
; Sequence 1109, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1109
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (80)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1109
Query Match      80.4%; Score 304; DB 9; Length 92;
Best Local Similarity 98.3%; Pred. No. 2.1e-30;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1 VLLRKEWTIGRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTDTSGTVINKLVV 60
Db      33 VLLRKEWTIGRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTDTSGTVINKLVV 92
|||||
```

```
RESULT 4
US-09-764-864-1110
; Sequence 1110, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1110
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1110
Query Match      72.8%; Score 275; DB 9; Length 99;
Best Local Similarity 90.3%; Pred. No. 1e-26;
Matches 56; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
Qy      12 RRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTDTSGTVINKLVVKKQTCPLQTGD 71
Db      9 RDKGPD--FPNKLVS GDHCRIIVVDEKSGQVLTDTSGTVINKLVVKKQTCPLQTGD 66
Qy      72 VI 73
Db      67 VI 68
|||||
RESULT 5
US-09-764-864-1529
; Sequence 1529, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1529
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1529
Query Match      72.8%; Score 275; DB 9; Length 99;
Best Local Similarity 90.3%; Pred. No. 1e-26;
Matches 56; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
Qy      12 RRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTDTSGTVINKLVVKKQTCPLQTGD 71
Db      9 RDKGPD--FPNKLVS GDHCRIIVVDEKSGQVLTDTSGTVINKLVVKKQTCPLQTGD 66
|||||
```

Qy 72 VI 73
Db 67 VI 68

US-09-740-627-11
; Sequence 11, Application US/09740627
; Patent No. US20020012964A1
; GENERAL INFORMATION:
; APPLICANT: Nakanishi, Makoto
; TITLE OF INVENTION: CELL CYCLE REGULATING FACTOR
; FILE REFERENCE: 06501-071001
; CURRENT APPLICATION NUMBER: US/09/740,627
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PCT/JP99/03350
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: JP 10/192467
; PRIOR FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae

US-09-740-627-11

Query Match 28.4%; Score 107.5; DB 9; Length 522;
Best Local Similarity 37.7%; Pred. No. 0.0001;
Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;

Qy 5 KREWTIGRRGCDLSPSPNKLVSQDHCRIIVVDEKSGQVLEDTSTGTVINKLVKKQT 64
Db 63 KKVWTFGRNPACDYHLGNISRLSNKHFQILLGE-DGNLLINDISTNGTWLNGQVKN 121

Qy 65 CPLQTGDVI 73
Db 122 QLLSQGDEI 130

US-10-081-119-16
; Sequence 16, Application US/10081119
; Publication No. US20030045491A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Jefferson, Anne B.
; APPLICANT: Chan, Vivien W.
; TITLE OF INVENTION: TTK in Diagnosis and as a Therapeutic
; FILE REFERENCE: 16932.002
; CURRENT APPLICATION NUMBER: US/10/081,119
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/289,813
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae

US-10-081-119-16

Query Match 28.4%; Score 107.5; DB 14; Length 821;
Best Local Similarity 37.7%; Pred. No. 0.00018;
Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;

Qy 5 KREWTIGRRGCDLSPSPNKLVSQDHCRIIVVDEKSGQVLEDTSTGTVINKLVKKQT 64
Db 63 KKVWTFGRNPACDYHLGNISRLSNKHFQILLGE-DGNLLINDISTNGTWLNGQVKN 121

Qy 65 CPLQTGDVI 73

Db 122 QLLSQGDEI 130

RESULT 8
US-10-032-585-7635
; Sequence 7635, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7635
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Candida albicans

US-10-032-585-7635

Query Match 26.2%; Score 99; DB 14; Length 699;
Best Local Similarity 33.3%; Pred. No. 0.0017;
Matches 23; Conservative 12; Mismatches 34; Indels 0; Gaps 0;

Qy 5 KREWTIGRRGCDLSPSPNKLVSQDHCRIIVVDEKSGQVLEDTSTGTVINKLVKKQT 64
Db 59 KQWYFGRDPSNDLQVASSRSISNKHQFQIWLNFNDKSLWIKDTNGTHLANSRLVKG 118

Qy 65 CPLQTGDVI 73
Db 119 YLLNQGDEI 127

RESULT 9
US-10-424-599-178873
; Sequence 178873, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 178873
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Glycine max

US-10-424-599-178873

Query Match 21.4%; Score 81; DB 12; Length 445;
Best Local Similarity 29.6%; Pred. No. 0.18;
Matches 24; Conservative 10; Mismatches 27; Indels 20; Gaps 2;

Qy 13 RRGCDLSPSPNKLVSQDHCRIIVVDEKSGQVLEDTSTGTVINKLV 59
Db 29 RRGCDLSPSPNKLVSQDHCRIIVVDEKSGQVLEDTSTGTVINKLV 59

Qy 60 VKKQT-----CPLQTGDVI 73
Db 89 AKKQVHLPNKETALENGDLV 109


```
RESULT 10
US-10-425-114-56100
; Sequence 56100, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56100
; LENGTH: 1082
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMLE01810018D04_FLI.psp
US-10-425-114-56100

Query Match      21.4%; Score 81; DB 12; Length 1082;
Best Local Similarity 33.8%; Pred. No. 0.55;
Matches 23; Conservative 12; Mismatches 27; Indels 6; Gaps 4;

QY 10 IGRRGCDLSPSPN-KLVSGDHCRIVVDEKSGQ--VTLEDTSSTGTV-INKLKVVKQTC 65
DB 1 IGSNRSC--NFFPLNDQTSIGNLKIKHTQGDGSAVAVLESMGSGVLVNGTHVKQNTSC 58
QY 66 PLOTGDVI 73
DB 59 VLNSGDEV 66

RESULT 11
US-09-867-550-1026
; Sequence 1026, Application US/09867550
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1026
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1026

Query Match      20.8%; Score 78.5; DB 9; Length 175;
Best Local Similarity 32.7%; Pred. No. 0.12;
Matches 18; Conservative 15; Mismatches 17; Indels 5; Gaps 2;

QY 24 KLVSGDHCRIVVDEKSGQVTLTDT-STSGTVINKLKVVK----KQTCPLQTGDVI 73
DB 49 KVLNRNALVWFDHKTGYLQDTKSSNGTFINSQRLSRGSESPPCILSGDII 103

RESULT 12
US-09-867-550-1026
; Sequence 1026, Application US/09867550
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1026
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1026

Query Match      20.8%; Score 78.5; DB 9; Length 175;
Best Local Similarity 32.7%; Pred. No. 0.12;
Matches 18; Conservative 15; Mismatches 17; Indels 5; Gaps 2;

QY 24 KLVSGDHCRIVVDEKSGQVTLTDT-STSGTVINKLKVVK----KQTCPLQTGDVI 73
DB 49 KVLNRNALVWFDHKTGYLQDTKSSNGTFINSQRLSRGSESPPCILSGDII 103
```

```
US-09-764-864-1111
; Sequence 1111, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1111
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (74)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (159)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (203)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (229)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1111

Query Match      20.8%; Score 78.5; DB 9; Length 244;
Best Local Similarity 32.7%; Pred. No. 0.18;
Matches 18; Conservative 15; Mismatches 17; Indels 5; Gaps 2;

QY 24 KLVSGDHCRIVVDEKSGQVTLTDT-STSGTVINKLKVVK----KQTCPLQTGDVI 73
DB 77 KVLNRNALVWFDHKTGYLQDTKSSNGTFINSQRLSRGSESPPCILSGDII 131

RESULT 13
US-09-740-627-10
; Sequence 10, Application US/09740627
; Patent No. US20020012964A1
; GENERAL INFORMATION:
; APPLICANT: Nakanishi, Makoto
; TITLE OF INVENTION: CELL CYCLE REGULATING FACTOR
; FILE REFERENCE: 06501-071001
; CURRENT APPLICATION NUMBER: US/09/740,627
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PCT/JP99/03350
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: JP 10/192467
; PRIOR FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-09-740-627-10

Query Match      20.5%; Score 77.5; DB 9; Length 460;
Best Local Similarity 30.8%; Pred. No. 0.52;
Matches 22; Conservative 9; Mismatches 34; Indels 7; Gaps 2;

QY 8 WTTGRRGCDLSPSPNKLVSQVTLTDTSTSGTVINKLKVVK-----VDEKSGQVTLTDTSTSGTVINKLKVVK 61
DB 60 WRFGRHKSCEVVL-NGPRVSNFHFYIYQGRNDSDESNVFLHDHSSNGTFLNFERLAK 118
QY 62 KQTCPLQTGDVI 73
DB 119 NSRTILSNGDEI 130
```

QY 59 VVKQTCTPQTGDVI 73
DB 154 VERNTYVDLHVGDVI 168

Search completed: May 7, 2004, 15:06:50
Job time : 12.5628 secs

RESULT 14
US-10-618-173-2
; Sequence 2, Application US/10618173
; Publication No. US20040018603A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: GUO, Zijian
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CHECKPOINT-ACTIVATING OLIGONUCLEOTIDES
; FILE REFERENCE: CITI350-1
; CURRENT APPLICATION NUMBER: US/10/618,173
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US/09/849,617
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/202,028
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-618-173-2

Query Match 20.2%; Score 76.5; DB 15; Length 517;
Best Local Similarity 31.9%; Pred. No. 0.81;
Matches 23; Conservative 9; Mismatches 29; Indels 11; Gaps 2;
QY 7 EWTIGRRRGCDLSF-----PSNKLVSGDHCRIVVDEKSGQ---VTLEDTSTSGTVIN 55
DB 84 EYVFGDRKKCDYTFDIPVLNQTDYKTSKRHFRIPOELGHHGHSRVANTIEDLSGNGTFVN 143
QY 56 KLVVKKQTCTPL 67
DB 144 KEIIGKRTLPL 155

RESULT 15
US-10-424-599-187236
; Sequence 187236, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J.
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187236
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(229)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140086C.1.pap
US-10-424-599-187236

Query Match 19.7%; Score 74.5; DB 12; Length 229;
Best Local Similarity 36.0%; Pred. No. 0.52;
Matches 27; Conservative 9; Mismatches 32; Indels 7; Gaps 4;
QY 2 LLKREWTIGRRGCD--LSFPNKLVSGDHCRIVVDEKSGQVLEDT-STSGTVINKLK 58
DB 98 VFKEGAYMFGRLDLCDFVLEHPT---ISRPHA-WQFKSGDAYLYDLGSTHGTFLNKQ 153

1	1995	74.5	1995	9	US-09-780-525-1	Sequence 1, Appl
2	1919.4	71.6	2448	16	US-09-108-260A-2073	Sequence 273, App
3	945	35.3	1311	16	US-09-764-864-22	Sequence 202, Appl
4	876.2	32.7	2186	16	US-10-094-749-160	Sequence 160, App
5	536.6	20.0	693	9	US-09-764-864-483	Sequence 483, App
6	394.4	14.7	476	10	US-09-918-995-2180	Sequence 2180, App
7	210.2	11.6	357	9	US-09-764-864-332	Sequence 322, App
8	273.6	10.2	449	10	US-09-918-995-13662	Sequence 13662, App
9	256.4	9.6	354	9	US-09-764-864-323	Sequence 323, App
10	256.4	9.6	354	9	US-09-764-864-742	Sequence 742, App
11	60	2.2	60	10	US-09-908-975-6220	Sequence 6220, App
12	51.4	1.95	1355	13	US-10-424-599-95539	Sequence 95539, App
13	51	1.9	383	10	US-09-814-353-18006	Sequence 18006, App
14	51	1.9	1411	13	US-10-424-599-127666	Sequence 127666, App

	Query Match	74.5%	Score 1995;	DB 9;	Length 1995;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1995;	Conservative	0;	Mismatches	0; Indels
					0; Gaps
QY	91	ATGAGCGGCCGAGGAGGACGAGTCGCGCCGCCGCGACGCCCTGGGGACGGCTCTTG	150		
DB	1	ATGAGCGGCCGCGAGGAGGACGAGTCGCGCGCGCGACGCCCTGGGGACGGCTCTTG	60		
QY	151	CGTCTGGGCGCGGANGGCGGCGACGCTCCCTTGAGGAGCGGGAGCTGGACCATC	210		
DB	61	CGTCTGGGCGCGAGAGGCGCGACGCTCCCTTGAGGAGCGGGAGCTGGACCATC	120		
	211	CGGCGGAGACGAGCTGGACCTTCTCTTCCGCCAGCAATAACTGCTCTTGGAGATCAC	270		

Db	887	ATAAATCACTCAGACATGCTGACGCCAAAGTCAGGCGTCTTTTCTGATGAAGAG	946
Qy	1298	GGAGTTGAGAGACCTGCTGGAGCTGTGACGTTTGACGAGTGTCTTCAGACATTAGCC	1357
Db	947	GGAGTTGAGAGACCTGCTGGAGCTGTGACGTTTGACGAGTGTCTTCAGACATTAGCC	1006
Qy	1358	AGCCATACGTCTGTGTCGGCAGTGTCTCTGAGTACAGAAGGCAGGCGCGCAGCCCTCCCC	1417
Db	1007	AGCCATACGTCTGTGTCGGCAGTGTCTCTGAGTACAGAAGGCAGGCGCGCAGCCCTCCCC	1066
Qy	1418	ACTGCCACACCCGAGGGCGAGCCAGAGACCCCAACAGGCCCTGGGGATGACCCCTCCA	1477
Db	1067	ACTGCCACACCCGAGGGCGAGCCAGAGACCCCAACAGGCCCTGGGGATGACCCCTCCA	1126
Qy	1478	CGTCCGTGACGCTGACGACAGCAGTCCAGGATTACGTGTGCCCTCTGCAAGGAAGCCACG	1537
Db	1127	CGTCCGTGACGCTGACGACAGCAGTCCAGGATTACGTGTGCCCTCTGCAAGGAAGCCACG	1186
Qy	1538	CCCTGTGCACCTGTCTCTTCAGGCCATGCCCCGACCGGAGACGGCGCGAGCAGGACC	1597
Db	1187	CCCTGTGCACCTGTCTCTTCAGGCCATGCCCCGACCGGAGACGGCGCGAGCAGGACC	1246
Qy	1598	CGGCTGTGCCCTCAGCAGTGTGGGTCTGCTCAGCCTTCTGCCACCTGTACTGGG	1657
Db	1247	CGGCTGTGCCCTCAGCAGTGTGGGTCTGCTCAGCCTTCTGCCACCTGTACTGGG	1306
Qy	1658	GTGTACCCGAGCCGGCTGCTACGGCTGCTGGCCCCGCTTTTGT	1701
Db	1307	GCTGCACCCGGACCGGCTGCTACGGCTGCTGGCCCCGCTTTTGTGTGTCGGCAGACAGT	1366
Qy	1702	-----GAGCTCAA	1709
Db	1367	GCTTTCCTCGCACATCATTTATAACAGGTAAACATTTGCCCTTCATCTCGGAGAGTCAA	1426
Qy	1710	CTTGGGTGACAAAGTGTCTGGAAGGGCTGTGTAACAACAAACAGCTACGAGTCTGACATCT	1769
Db	1427	CTTGGGTGACAAAGTGTCTGGAAGGGCTGTGTAACAACAAACAGCTACGAGTCTGACATCT	1486
Qy	1770	GAAGAAATTACTGGCMAACAGAGGTTTGACATGGAAGAAACATGTTCACCGAGAGCCCTCGT	1829
Db	1487	GAAGAAATTACTGGCMAACAGAGGTTTGACATGGAAGAAACATGTTCACCGAGAGCCCTCGT	1546
Qy	1830	GGCTCTCCAGCGGGGAGTGTTCCTGCTGTCAATTACAGAGTCAACGGGAGACACCGTTCT	1889
Db	1547	GGCTCTCCAGCGGGGAGTGTTCCTGCTGTCAATTACAGAGTCAACGGGAGACACCGTTCT	1606
Qy	1890	GTGTTACTGCTGTGGCTCGGAGCTTCGCTGAGCTGACCTATCAGTATCGGCAGAACAT	1949
Db	1607	GTGTTACTGCTGTGGCTCGGAGCTTCGCTGAGCTGACCTATCAGTATCGGCAGAACAT	1666
Qy	1950	TCTGCTTCCGAGTGTCCCAAGTGGCGGTAAACATCCCTCCTGACTGCTACTTGGGGCCGTAA	2009
Db	1667	TCTGCTTCCGAGTGTCCCAAGTGGCGGTAAACATCCCTCCTGACTGCTACTTGGGGCCGTAA	1726
Qy	2010	CTCCGCACTCAGGTGAAGCTCAACACGCCATGAAATTCATCATATCTGTGAAACAGAC	2069
Db	1727	CTCCGCACTCAGGTGAAGCTCAACACGCCATGAAATTCATCATATCTGTGAAACAGAC	1786
Qy	2070	AAGGTTCAAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTTCAGCACTGGAGGTGAAGAG	2129
Db	1787	AAGGTTCAAAAACCTAAGCATCCAGAGGCCCTGAGCAGCTTTTCAGCACTGGAGGTGAAGAG	1846
Qy	2130	AGCGTGTTTTTAAAAATACAGACAAGACAGTCAAGGTGTTTTTCAGAGCCCCCTGAGGGA	2189
Db	1847	AGCGTGTTTTTAAAAATACAGACAAGACAGTCAAGGTGTTTTTCAGAGCCCCCTGAGGGA	1906
Qy	2190	AGGAGCGCAGGCTCCGACAGTGTCTGGGGTGACTCTTCTGTGGAGCTTTTATACCCCT	2249
Db	1907	AGGAGCGCAGGCTCCGACAGTGTCTGGGGTGACTCTTCTGTGGAGCTTTTATACCCCT	1966
Qy	2250	CTGAGTGAGACCTCCCCAGAGGCCCGGGGGCCGACGCCGCCCTCTGTGTGAGCGCTGG	2309

LOCATION: (623)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (640)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (670)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (687)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (690)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-483

Query Match 20.0%; Score 536.6; DB 9; Length 693;
Best Local Similarity 92.5%; Pred. No. 1.6e-138;
Matches 608; Conservative 7; Mismatches 13; Indels 29; Gaps 4;
QY 1044 GGCCTGATGACACAGTTCTGCGGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCT 1103
DB 14 GGCCTGATGACACAGTTCTGCGGGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCT 73
QY 1104 GTGTCTTACCTGCCCTGTCCTGCGGGATCTGTAAACACATCTCTCAACACCT 1163
DB 74 GTGTCTTACCTGCCCTGTCCTGCGGGATCTGTAAACACATCTCTCAACACCT 133
QY 1164 CTTGAAGCATACCTCATCTACGATCCAGACAGAGTCCGAGTGAAGAGATGTGCAAG 1223
DB 134 CTTGAAGCATACCTCATCTACGATCCAGACAGAGTCCGAGTGAAGAGATGTGCAAG 193
QY 1224 TATGATGCCAGGAATAAATCACTCAAGACATGCTGAGCCCAAGTCAAGCGTCTTT 1283
DB 194 TATGATGCCAGGAATAAATCACTCAAGACATGCTGAGCCCAAGTCAAGCGTCTTT 253
QY 1284 TTCTGATGAAGAGGAGTTTCAAGAGCTGCTGAGCTGTGACAGTGTGACAGTGTG 1343
DB 254 TTCTGATGAAGAGGAGTTTCAAGAGCTGCTGAGCTGTGACAGTGTGACAGTGTG 313
QY 1344 CTCAGACATTACCCAGCCATACGTCTGTGCGGCACTGTCTGTGATGACAGAGCGGC 1403
DB 314 CTCAGACATTACCCAGCCATACGTCTGTGCGGCACTGTCTGTGATGACAGAGCGGC 373
QY 1404 GCGGAGCTTCCCTGCTGCGGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1463
DB 374 GCGGAGCTTCCCTGCTGCGGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 433
QY 1464 GGATGACACCTCCAGCTCCGCTGACGACAGCAGTCCAGGATACGTGTGCGCTCT 1523
DB 434 GGATGACACCTCCAGCTCCGCTGACGACAGCAGTCCAGGATACGTGTGCGCTCT 489
QY 1524 GCAAGAGCCAGCCGCTGTGACCTGCTGCTTCCAGCCCATGCGGAGGAGGAGGAGG 1583
DB 490 GCAAGAGCCAGCCGCTGTGACCTGCTGCTTCCAGCCCATGCGGAGGAGGAGGAGG 548
QY 1584 GCGGAGCAGGAGCCGCTGTGCGGCTCCAGCAGTGTGCGGCTGCTGCTGCTGCTG 1643
DB 549 -CGGAGCAGGAGCCGCTGTGCGGCTCCAGCAGTGTGCGGCTGCTGCTGCTGCTG 584
QY 1644 CCACTGTACTGGGGTGACACCGGAGCCGCTGTGCTGCTGCTGCTGCTGCTGCTG 1700
DB 585 CCACCTGTACTGGGGTGACACCGGAGCCGCTGTGCTGCTGCTGCTGCTGCTGCTG 641

RESULT 6
US-09-918-995-2180
Sequence 2180, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED FROM VARIOUS CDNA LIBRARIES

FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2180
LENGTH: 476
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(476)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2180

Query Match 14.7%; Score 394.4; DB 10; Length 476;
Best Local Similarity 96.7%; Pred. No. 5.5e-99;
Matches 413; Conservative 0; Mismatches 12; Indels 2; Gaps 1;
QY 2046 ATTCAATCATATCTGTGAACACAGAGGTTCAAAAATCAAGCATCCAGAGCCCTGAGCA 2105
DB 52 ATTCAATCATATCTGTGAACACAGAGGTTCAAAAATCAAGCATCCAGAGCCCTGAGCA 111
QY 2106 GCTTTTCAGCACTGGAGGTGAAGAGAGCGTGTGTTTAAATACAGACACAGCAAGCGTCAAG 2165
DB 112 GCTTTTCAGCACTGGAGGTGAAGAGAGCGTGTGTTTAAATACAGACACAGCAAGCGTCAAG 171
QY 2166 GTGTTTTTCAGCCCCCTGAGGGAAGGAGCGAGGCTCTCCGACAGGTGCTCTGGGGTGA 2225
DB 172 GTGTTTTTCAGCCCCCTGAGGGAAGGAGCGAGGCTCTCCGACAGGTGCTCTGGGGTGA 231
QY 2226 CTCTTCTGTGGAGCTTTTACCTCTGAGTGAGACCTCCCGAGAGCCCGGGGGCGGCA 2285
DB 232 CTCTTCTGTGGAGCTTTTACCTCTGAGTGAGACCTCCCGAGAGCCCGGGGGCGGCA 289
QY 2286 GCCCGCCCTCTCTGAGGAGGCTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2345
DB 290 GACCGCCCTCTCTGAGGAGGCTGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 349
QY 2346 CCTTCTGTAAATGCGGCGCTCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2405
DB 350 CCTTCTGTAAATGCGGCGCTCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 409
QY 2406 CCGAACTACAGTTAAAGCAGAGTCTGTTTTTCAAGGAAAGTTTCAAGGAGAGGAGGCA 2465
DB 410 CCGAACTACAGTTAAAGCAGAGTCTGTTTTTCAAGGAAAGTTTCAAGGAGAGGAGGCA 469
QY 2466 GTTTATC 2472
DB 470 GTTTATC 476

RESULT 7
US-09-764-864-322
Sequence 322, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PZ23
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 322
LENGTH: 357
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE

; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 742
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (289)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (313)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-742

Query Match 9.6%; Score 256.4; DB 9; Length 354;
Best Local Similarity 93.3%; Pred. No. 1.3e-60;
Matches 266; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
QY 221 GAGGTGGACCTTCCTCCCGCAATAACTGGTCTCTGGAGATCACTAGAAATG 280
DB |||||
QY 281 TAGTGGATGAAAAATCAGGTGACGTGACACTGGAAGATACACGACAGTGGAAACAGTGA 340
DB 87 TAGTGGATGAAAAATCAGGTGACGTGACACTGGAAGATACACGACAGTGGAAACAGTGA 146
QY 341 TTAACAGCTGAAGGTTGTTAAGAGCAGACATGCCCTTTACAGATCGGGATGCTACT 400
DB 147 TTAACAGCTGAAGGTTGTTAAGAGCAGACATGCCCTTTACAGATCGGGATGCTACT 206
QY 401 ACTTGGTGTACAGGAAGAATGAACCGGAACACACGTCGTCATACCTCTATGATCTTTAA 460
DB 207 ACTTGGTGTACAGGAAGAATGAACCGGAACACACGTCGTCATACCTCTATGATCTTTAA 266
QY 461 GTGAAGACGACGATGACACAGATCCTTTGAAGCTACAAGG 505
DB 267 GTGAAGACGACGATGACACAGATCCTTTGATACCTCAGG 311

RESULT 11
US-09-908-975-6220
; Sequence 6220, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6220
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-6220

Query Match 2.2%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1843 GGAGTGTCTTCTGCTGTGATTACAGATCAGGGAGACACCGTCTCTGTGTTACTGCTGT 1902
DB 1 GGAGTGTCTTCTGCTGTGATTACAGATCAGGGAGACACCGTCTCTGTGTTACTGCTGT 60
RESULT 12
US-10-424-599-95539
; Sequence 95539, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95539
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(1255)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57283C.1
US-10-424-599-95539

Query Match 1.9%; Score 51.4; DB 13; Length 1255;
Best Local Similarity 57.0%; Pred. No. 0.0047;
Matches 94; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 2515 ACAGGCGTACACAAATATCTGCTGCGGGAACACACAGCATTTTACTATTTTATT 2574
DB 858 ACATGGTGGAACCTTTAACTTGGCGTTGTAATAAGAACACCTTTGATCTCTATTTTC 917
QY 2575 TTAATAGTTTGGTGTCTTATCTTCTAATAAGATTTAAATGTCAAACTGTAGCAAAAT 2634
DB 918 TTGCTAGTAATAGGATGTTTACTTCTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 977
QY 2635 AATATAATTTAATTTTACAAATTTGACAAAAAATAAAAAAAAAAAAAA 2679
DB 978 AA 1022

RESULT 13
US-09-814-353-18006
; Sequence 18006, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-0068
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25

```

/ APPLICANT: Tao, Nengbing
/
/ APPLICANT: Byatt, John C.
/
/ APPLICANT: Mathialagan, Nagappan
/
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/
/ TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
/
/ FILE REFERENCE: 16511.006/37-21(10298)C
/
/ CURRENT APPLICATION NUMBER: US/09/960,352
/
/ CURRENT FILING DATE: 2001-09-24
/
/ NUMBER OF SEQ ID NOS: 15112
/
/ SEQ ID NO 3525
/
/ LENGTH: 434
/
/ TYPE: DNA
/
/ ORGANISM: Bos taurus
/
/ OTHER INFORMATION: Clone ID: 16-BOVMS1-023-Q1-E1-D11
/
/ US-09-960-352-3525

```

[illegible]

Search completed: May 15, 2004, 00:10:39
Job time : 1598 secs

```

/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 127666
/ LENGTH: 411
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MBT3847_86287C.1
/ US-10-424-599-127666

```

	Query Match	1.9%	Score 51;	DB 13;	Length 411;
	Best Local Similarity	63.4%;	Pred. No. 0.0028;		
	Matches	78;	Conservative	0;	Mismatches 45;
				Indels	0;
				Gaps	0;
Qy	2557	TTTTATCTATTTTTTTTAAATAGTGTGGTCTTATCTTCTAAATGAATTTAAATGTC	2616		
Ddb	216	TCATATCAATTTTGTGTGGTGTGGTGTATTTGTTTGGTGTGTTTAACTTTA	157		
Qy	2617	ACAACTGTGACACAAATATATAATTTATAATTTACAAATTCACAAAAAATAAAAAA	2676		
Db	156	AAATAATTTTAAAAAATAAAAAATAAAAAAATACCAAAAAAATAAAAAAATAAAAAA	97		
Qy	2677	AAA	2679		
Db	96	AAA	94		

RESULT 15
US-09-960-352-3525/c
; Sequence 3525, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 12:12:53 ; Search time 10037 Seconds
(without alignments)
7970.587 Million cell updates/sec

Title: US-10-048-046-1
Perfect score: 2679
Sequence: 1 aagaattcgccagggccg.....acaaaaaaaaaaaaaaaaa 2679

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_estci.*
9: gb_estli.*
10: gb_estc2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_lam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1464.6	54.7	3146	11	AK077629 Mus muscu
2	1366.2	51.0	2876	11	AK048930 Mus muscu
3	919	34.3	1201	9	AL539798
4	911.8	34.0	1201	9	AL539797 AL539797

5	822.2	30.7	876	13	BQ876252
6	774.8	28.9	921	13	BUI176559
7	765.4	28.6	870	13	BQ215731
8	758	28.3	992	13	BQ073354
9	757.6	28.3	877	14	CD243458
10	740.6	27.6	910	13	BUI148471
11	727.4	27.2	840	12	BQ395714
12	726.8	27.1	953	13	BQ396072
13	714	26.7	938	13	BQ924736
14	708.4	26.4	824	12	BQ831289
15	706.2	26.4	944	13	EX409790
16	700	26.1	909	13	EX409791
17	694.4	25.9	729	14	CA916737
18	694.4	25.9	1201	9	AL561570
19	680	25.4	1116	9	AL528043
20	674.8	25.2	897	10	BF971929
21	661	24.7	940	13	BQ934867
22	652.4	24.4	759	12	BI769973
23	650.4	24.3	916	12	BM459647
24	639.6	23.9	666	10	BE407732
25	637.4	23.8	934	10	BF971838
26	637.2	23.8	1092	12	BM541730
27	628	23.4	687	9	AUI138310
28	626.8	23.4	690	10	BE280165
29	626.2	23.4	938	13	BQ887242
30	624.8	23.3	1022	12	BM558904
31	624.8	23.3	1096	12	EG335362
32	617	23.0	722	9	AUI135780
33	611	22.8	680	9	AUI134937
34	602.6	22.5	636	10	BE407797
35	588	21.9	635	12	EG481936
36	587.6	21.9	824	13	BQ572677
37	586.6	21.9	817	12	BM944161
38	583	21.8	824	9	AUI132825
39	578.2	21.6	712	10	BE257482
40	567.8	21.2	650	10	AW957709
41	567.4	21.2	811	14	CB989147
42	566.4	21.1	1143	12	BM542545
43	559.6	20.9	806	14	CB961597
44	536.2	20.0	982	13	CB943529
45	516	19.3	782	14	CF410232

ALIGNMENTS

RESULT 1
AK077629
LOCUS AK077629 3146 bp mRNA linear HTC 18-SEP-2003
DEFINITION Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730494M20 product:similar to CELL CYCLE CHECKPOINT PROTEIN CHFR [Homo sapiens], full insert sequence.
ACCESSION AK077629
VERSION AK077629.1 GI:26346521
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Taehiro,H., Iton,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Iihikawa,T., Oawa,K., Tanaka,T., Matsumura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
REFERENCE

AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
REFERENCE
AUTHORS

6 (bases 1 to 3146)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaki-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toyota,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

JOURNAL
REFERENCE
AUTHORS

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES
source

Location/Qualifiers
1. 3146
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM,DB:5730484M20"
/db_xref="MGI:2393804"
/db_xref="taxon:10090"
/clone="5730484M20"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="8 days embryo"
99. -2080
/note="unnamed protein product; putative
similar to CELL CYCLE CHECKPOINT PROTEIN CHFR [Homo
sapiens] (SPTK|Q9NRT4, evidence: FASTV, 81.7%id,
100%length, match=1989)"
/codon_start=1
/protein_id="BAC36912.1"
/db_xref="GI:26346522"

CDS

1. 3146
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM,DB:5730484M20"
/db_xref="MGI:2393804"
/db_xref="taxon:10090"
/clone="5730484M20"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="8 days embryo"
99. -2080
/note="unnamed protein product; putative
similar to CELL CYCLE CHECKPOINT PROTEIN CHFR [Homo
sapiens] (SPTK|Q9NRT4, evidence: FASTV, 81.7%id,
100%length, match=1989)"
/codon_start=1
/protein_id="BAC36912.1"
/db_xref="GI:26346522"

/translation="MELHGEQPPPPQEPWAGPWRLLRLGAREDEPQILLWKREWTIGRR
CGLSPFNKLVSGDHCKLVDSIGEVLTSTNGTVINKLVQVKKQYIPLHSGDI
IYIVRKNEPHEVAYLYESLQSLQTSLEANKENFMHTVKQCSGGQGDQVVP
LLSPMAQTLSEPPSTSDLPSTASTSTPELTSPAKKSSSGSPGNTSISPKGR
SSLVANGELSSLPVQDEASPSLSEKSHSELEPAKKMKGDGDELDTNLLVSGQ
RGNACTSSBVDKASVPCDKMBELTLCIIQDLDHCVSLQPCMTFFCAACYSQWNER
SLLCPTCRCPVERICKNHLINLVAVLIQHPKSRSEEDVSMADARKITQDMLQPK
VRRSFDEEGSSDLLELSDVPSSESDISQPVIVCQCPYERQAVQSPICVPSESEL
GATLALGGEAPSTASLPTAPYMCPLQSGHAICTCCFPQMPDRAREQSDPRVAPQ
CAVLQPFCHLYMGCTKTCFCLAPFCELNLDGKLDGLVANNNNYESDILKYLATR
GLTVTSRPPDPCYWGRCRTQVYAHAMKFNHICEQTRFNK"

polyA_signal

/notes="putative"

polyA_site

/notes="putative"

ORIGIN

Query Match 54.7%; Score 1464.6; DB 11; Length 3146;
Best Local Similarity 74.5%; Pred. No. 2.3e-233;
Matches 2007; Conservative 0; Mismatches 634; Indels 52; Gaps 11;
QY 18 CCGCAATGCTCTTTGACAGCGCGCGCGCG- - - - -CAGCGGTTCCGGGTTCCGGCGGG 71
DB 10 CCGAGATGCTCTTTGACAGCGGTGGCGGACTGTGAGGCCGGTTCCGGSATCCGCGCAG 69
QY 72 GCGCGGAGATGTAATCCCGATGAGCGCGCGGAGGAGCAAGCAGTCCGCCCGCGCGCA 131
DB 70 GCGGTGGGATGTGAGTCTTATGAGCTACACGCGGAGAGCAGCGCGCGCGCGCAGGA 129
QY 132 GCCCTGGGGACGCTCTCTGCTGGCGCGGAGGAGGCGGCGCGCAGCTCCTCTCTGAG 191
DB 130 ACCCTGGGGAGGCTTCTGCTAGCGCGCAGGAGGAGCAACCCAGATCTCTTTTG 189
QY 192 GAGCGGAGATGACCATCGGCGGAGAGAGAGTTGCGACCTTTCTTCCCGCAGCAATAA 251
DB 190 GAAACCGAGTGGACCATCGGAGGAGGAGGAGGCTGTGACCTCTCTTCCCGCAGCAATA 249
QY 252 ACTGGTCTCTGGAGATCATCTGTAGAAATTTAGTGGATGAAATCAGGTGAGTGCACACT 311
DB 250 ACTGGTCTCTGGAGATCATCTGTAAACTTACAGTGGATGAATATCTGGTGGTGCACACT 309
QY 312 GGAAGTACCCAGCAGCAGTGGAAACAGTGAATTAACAGCTGAAGTTGTTAAGAACGAGAC 371
DB 310 GGAAGACACAGCAGCAATGGAAACAGTGAATCAATAGCTGCAAGTTGTTAAGAACGAGAC 369
QY 372 ATGCCCTTTACAGATGGGGATGTCTATCTACTTGGTGTACAGAAAGAAATGAACCGCAACA 431
DB 370 TTACCTTTACATAGCGGGATATCATCTATTGGTGTACAGAAAGATGACCCAGACA 429
QY 432 CAACGTGGCATACCTCTATGAATCTTTAAGTGAAGAAAGAGCATGACACAGAAATCTCT 491
DB 430 CAATGTGGCATACCTCTATGAATCTTTAAGTGGCAACAGAGCTTAACCTCAGACTCCCT 489
QY 492 TGAAGCTTAACAGAAATGTGTTCATGGGACCAAGATACCTCAGGTGCGAGGTGCAGG 551
DB 490 TGAAGCAATAAGAAATATGTTCCATGTGACCAAGATTGCT- - - - -CAGGTCCAGG 543
QY 552 GCGAGGGCGGATCCCGGGTCCCTCCGTCGTCGCGCGCGCAGTCAAGTGTGCTTTGAGGA 611
DB 544 GCAGGGTGAATGATCCCGAGGTTCCACTATTTGTCCACCCATGGCTCAGACATGCTTAGAGA 603
QY 612 ACCACAGCCATCAACATCCAGCTCAGACCTCTTCCCGCAGCAGCTCGGCTCTTTCACCGGA 671
DB 604 ACCACAGCCATCAACATCCAGCTCAGACCTCTCTCCCGCAGCAGCTCTACCTCTTCTACGGA 663
QY 672 GCC- - - - -TTCTCTGCGAGGCGGAGGCTTCTTCCAGTGTGGGTCTGGGGTGGTGG 725
DB 664 GCCAGAGCTGACCTCTGCGAGGCAAAAGCATTTCTTCTAGCTCTGAGCTGGGAAACACAAG 723
QY 726 CATCTCCCTAAAGGAAGTGTCCCTCTGTGGCAAGTGTGATGAAGTCTCCAGCTTTGCTCTC 785
DB 724 CATCTCCCAAGGAGCGAGTTCACTTGTTCGAAATGCGGAATCTCTTAGCCTTTCTCC 783

QY 786 AGCTCTCCAGACAGAAAGACTGCGTCTTTCGTCGTTGGAAACCCAGGATCAGAGGA 845
Db 784 AGTTTTCCAGACAAAGA---AGCATCTCTTTTCTTGTGCGAAAGTAAAGACCATGAGA 840
QY 846 TTGAGAGCCGTCAGACAGAAATAGAGAGATGGGACCTTGACCTGAACGGCGAGTT 905
Db 841 ATTGGAGCCTGCCAAAAAAGATGAAAGAGATGGGAACTTGACACGAACCTCCAGTT 900
QY 906 GTTGGTCGACAAACCCGCTAGAAATGCCAAACCGTCCACGAGGAGCTCAGAGCAGCGC 965
Db 901 ATTAGTTTCAGGCGAGCTGGAAATGCCAAACCTCAAGTGAAGATGTCAAAAGATGCCTC 960
QY 966 TGGAGACGACAGAAATGGAGAGAGCTGACATCATCTGCGAGGACTGCTGCA 1025
Db 961 TGTGAAGCCAGACAGATGAGAGAGACATAACCTGTATCATCTGCCAGAGACTTCTGCA 1020
QY 1026 CGACTGCGTGAAGTTTCAGCCCTGATGACACAGTTCTCGCGGCTTGTACTCGGGCTG 1085
Db 1021 CGATTGTGTGAGTTTCAGCCCTTGATGACACATTTTGTGGGCTTGTACTCTGGTTG 1080
QY 1086 GATGGAGCGTCTGCTGTGTCCTACCTGCGGCTGTCGCGTGGAGCGATCTGTAATAA 1145
Db 1081 GATGGAGCGTTCATCTGTGCGCTTACCTGCCGATGTCAGTGGAGCGAATTGCAAAA 1140
QY 1146 CCACATCTCAACAACTCGTGAAGCATACCTCATCCAGCATCCAGACAGAGTGGCAG 1205
Db 1141 CCACATCTCAACAACTAGTGAAGCATACCTTATCCAGCACCAGATAAAGTGGCAG 1200
QY 1206 TGAAGAGATGTGCAAGATGATGACAGGAATAAATCATCTCAAGAGATGCTGAGCTGC 1325
Db 1201 TGAAGAGATGTGAGAAGTATGATGCAAGGAATAAATCATCTCAAGATATGCTGCAAC 1260
QY 1266 CAAAGTCAGCGCTCTTTCTCATGAGAGGAGTTCAGAGAGCTCTGCGAGCTGTC 1325
Db 1261 CAAAGTCAGAGGCTCTTCTCATGAGAGGAGTTCAGAGAGCTCTGAGAGCTGTC 1320
QY 1326 AGACGTTGACAGTGAAGTCTCAGACATATGACCCAGCCATAGTCTGCTGCGGAGTCTC 1385
Db 1321 TGATGTCGACAGTGAATCTCTCAGATATCAGTACGCCATATATGTCGAGACAGTGTCC 1380
QY 1386 TGAGTACAGAGCAGCGCGGAGCTCCCTCCCTAGCAGCACCAGGAGGCGGAGCCAGG 1445
Db 1381 TGAATACAGAGCAAGCGGTGAGTCTCTTCTTCCGAGTCCAGAGAGTGAAGTGGG 1440
QY 1446 AGCCACAGCGCCTGGGGAGTGCACCTCCAGTCCGTGAGCTGACGACAGCAGTCCA 1505
Db 1441 AGCTACACTGGCCCTTGGTGGGAGGACCTTCAACATCTGCCAGCTTGCCAAACAGCCC 1500
QY 1506 GGATTAAGTGTGCGCTCTGCAAGAGACCGAGCCCTGTGCACTGCTGCTTCCAGGCCAT 1565
Db 1501 GGATTAAGTGTGCGCTCTTCAAGAGAGCCATGCAATATGCACTGCTGCTTCCAGCCAT 1560
QY 1566 GCCGACCGGAGCGAGCGGAGCAGGAGCAGGACCGCGTGTGCGCCCTCAGCAGTGTGCGGT 1625
Db 1561 GCCTGACCGGAGAGCTGAACGGGAGCAGGATCCCGGCTGCGCCCTCAGCAGTGTGCGGT 1620
QY 1626 CTGCTCGAGCTTTCTGCCACTGTACTGGGGTGCACCCGACCGGCTGTCTACGGCTG 1685
Db 1621 GTGCTCGAGCCCTTCTGCCACTGTACTGGGGTGCACGAGGACTGGCTGCTTTGCTG 1680
QY 1686 CCTGGCCCGTTTGTGAGTCAACCTGGGTGCAAGTGTCTGACGCGGTGCTGAACAA 1745
Db 1681 CTTGGCCCGATCTGTGAGCTCAACCTGGGGAGCAGTGTGATGGAGTGTGAGAACAA 1740
QY 1746 CAACAGCTACAGTACAGATCTTGAAGAATTAACCTGGCAACAGAGGTTTGAATGGAA 1805
Db 1741 TAACAATTAATCGGACATCTTGAAGAATTAACCTGGCAACAGGAGTCTGACATGGAA 1800
QY 1806 AAACATGTTGACAGAGCTCTGCGCTCCAGCGGGAGTGTCTGCTGCTCATTA 1865
Db 1801 AAGTGTGTGACAGAGTCTCTGCTCTGCGAGGAGGTATTTATGCTGCTCATTA 1860

QY 1866 CAGAGTCACGGGAGACACCGTTCTGTCTTACTGCTGTGGCTCGCAGCTTCCGTGAGCT 1925
Db 1861 CAGAACTACTGGAAATACTGTGCTGTGTTACTGCTGTGGTCTCGTACGTTCCGAGAGCT 1920
QY 1926 GACCTATCAGTATCGGAGAACATTTCTGTTTCGAGTTGCGAGTGCCTGTAACATCCG 1985
Db 1921 GACCTACAGTATCGTCAGAACATTTCTGTTTCGAGTTGCGAGTGCCTGTAACATCCG 1980
QY 1986 TCCTGACTGCTACTGCGGCGGTAACTGCGCACTCAGTGTAAAGCTCACACGCGCATGAA 2045
Db 1981 TCCTGACTGCTACTGCGGCGGTAACTGCGCACTCAGTGTAAAGCTCACATGCAATGAA 2040
QY 2046 ATTCAATCATATCTGTGAACAGACAGAGTTTCAAAAATAAGCATCCAGAGCCCTGAGCA 2105
Db 2041 ATTCAATCATATCTGTGAGCAAAACAAGTTTCAAAAATAAGCATCCAGAGCACGTGA 2100
QY 2106 GCTTTCAGCATGAGAGTGAAGAGAGCGTGTGTTTTTAAATATACAGAGACAGAGTCAAG 2165
Db 2101 AACAGTTACAGAGCA-CAGGAGGCAATGTTTTTAAAGGCCCCATACAAACATTTGAG 2159
QY 2166 GTGTTTTTACAGAGTTCAGGAGGAGGAGCGAGGCTCTCCGACAGGTGCTCTGGGGTGA 2225
Db 2160 GGCAATTTTACAGTTTCCCTACGCACTGCGGGTCTCTGGTCAGGCACTGGGGTGTGG 2219
QY 2226 CTCTTCTGTGAGAGCTTTTACCTCTGAGTGAAGACCTCCCGAGAGCCCGGGGGCGCA 2285
Db 2220 ATTTTCCCTTTGTGAGAGCTCATCTGTCAGTGACACACACACCGCTGCCAGTAAC-CT 2278
QY 2286 GCCGCGCTCTGCTGAGCGCTGGGAGGCTCGTGTGGTGGCATCAGCAGCAGAG- 2339
Db 2279 GGGAGGCTCTCTGCCACACACAGCGCTTCTAGTCTGTTAAGAGGATCAGTAGACTGAGTT 2338
QY 2340 -----ACGAAGCCCTTCTGTAAATCGCGGCGCTCCGAGAGTCTGTTTTTCAAGAAAAG 2446
Db 2339 CTCCCTCCCTCCCACTGGACCGCTCTCCCTCTGCTGTGAGGAGTGAAGTGCAGGT 2398
QY 2387 TTGCTCTTTTCTACATTTTCCGAAACTACAGTTAAAGCAGAGTCTGTTTTTCAAGAAAAG 2446
Db 2399 TTGTAATTTTGTATCTTTT-----CTACCTTGTACTGTTAAATGGGAAG 2452
QY 2447 TTTCAGGAGAGAGGCGCAAGTTTATCAAAAACATGTTTTCAGGAG--AAGGAGCATTAAG 2504
Db 2453 CTCATAGAAAAGTTGTCAGCTGTGTGAGCGCATTTCTCCGGAAGGAAGAGCATACG 2512
QY 2505 TTTACAGCTACAGAGCT-ACACATATCTGCTGTGGGAGAAACACACATTTTATC 2563
Db 2513 TTTACAGCTCCAGGGGTACATAGATAGTCTCTCTCTGGAAGACACAGCCCTTTGTA 2572
QY 2564 TATTTTTTATTTTATAGTTTGTGCTTA-TCTTCTAATAAGATTTAAATGTCAAAAC 2622
Db 2573 TATTTTTTATTTAATAAGTTTGTGCTTAGTCTTGTAAATAAGATTTAAATATTTAAAC 2632
QY 2623 TGTAGCAAAATATATATTTTATTTTAAATTTGACAAATTTGAGGAATA 2675
Db 2633 TGTAGCAAAATATATATTTTATTTTAAATTTGACAAATTTGAGGAATA 2685

RESULT 2
AK048930 2876 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length
DEFINITION enriched library, clone: C23082M18 product: similar to CELL CYCLE
CHECKPOINT PROTEIN CHFR [Homo sapiens], full insert sequence.
ACCESSION AK048930.1 GI:26093209
VERSION AK048930.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE

Qy	1107	TCCTACCTCGCGGTGTCCTCCGTCGAGCGGATCTCTGTAAAAACCAATCTCTCAACAACTCTGCT	1166
Db	1281	CCCTACCTGCGGATGTCCAGTCGGAGCGGATTTTGCAAAACCAATCTCTGAAACCACTTAGT	1340
Qy	1167	GGAGGCAATCCTCATCCAGCATCCAGACAAGAGTCGCAGGTGAAGAAGATGTGCAAAAGTAT	1226
Db	1341	GGAGGCAATCCTTATCCAGCAACCAGATAAAGTCCGAGTGGAGAGATGTGAGAGTAT	1400
Qy	1227	GGATGCCAGGAATAAATCACTCAAGACATGCTTGCAGCCCCAAAGTCAAGCGGTCTTTTTC	1286
Db	1401	GGATGCAAGGAATAAATCACTCAAGATATGCTTGCACCCCAAAAGTCAGGAGTCTTTCTC	1460
Qy	1287	TGATGAAGAGGAGTTTCAGAGACCTGCTGGAGCTGTGCAGCGTTGCACAGTCAGTCCCTC	1346
Db	1461	TGATGAAGAGGAGTTTCAGAGACCTGCTTAGAGCTGTCTGATGTGCACAGTGAATCCTC	1520
Qy	1347	AGACATTAGCCAGCATPACGTGTGTGCCCGCAGTGTCTTGCAGTACAGAGGAGGAGCGGC	1406
Db	1521	AGATATCAGTCAGGCATATCATTTGTCTGCACAGTGTCTTGAATACAGAAGGCAAGCGGT	1580
Qy	1407	GCAGCTCCCACTGCCAGCACCCAGGCGGAGCCAGGAGCCCCACAGCCCTGGGGGA	1466
Db	1581	GCAGTCTCTTCTTGGCCAGTCCCAAGAGATGAGCTGGGAGCTACACTGCGCCCTTGGTGG	1640
Qy	1467	TGCACCTTCCAGTCGCTCAGCCCTGACGACAGAGATCCAGGATTACGTGTGCCCTCTTGCA	1526
Db	1641	GGAGGCACCTTCAACATCTCCAGCTTGCACACAGCCCCGAGTTTACATGTGCCCTCTTCA	1700
Qy	1527	AGGAAGCAACGCTCTGTGCACTGTCTGCTTCCAGCCCATGCCCCGACCGGAGAGCGAGCG	1586
Db	1701	AGGAAGCATGCCATATGCACTGCTGCTTCCAGCCTATGCTTGACCGGAGAGCTGAACG	1760
Qy	1587	CGAGCAGGACCCGGGTGTCGCCCTCAGCAGTGTGCGGTGTGCTGCTGCAGAGCCTTTTCGCCA	1646
Db	1761	GGAGCAGATCCCGCGTTCGCCCTCAGCAGTGTGCGGTGTGCTGCAGAGCCTTTTCGCCA	1820
Qy	1647	CCTGTACTGGGGTGCACCCGAGCCGCTGCTACGGTGTGCTGGCCCGCTTTTGTGAGCT	1706
Db	1821	CCTGTACTGGGGTGCACAGGACTGSGCTGCTTTTGGTGTGCTTGGCCCATCTCTGTGAGCT	1880
Qy	1707	CAACCTGGGTGACAAAGTGTCTGGACGCGCTGCTGAAACAACAAGCTTACCAGTCAGACAT	1766
Db	1881	CAACCTGGGGACAAGTGTCTTGGATGGAGTGTGACAAATTAACATTAATGAATGGACAT	1940
Qy	1767	CCTGAAGAAATTAACCTGGCAACAGAGGTTTGACATGGAATAACATGTGACCGAGAGCTT	1826
Db	1941	CCTGAAGAAATTAACCTGGCAACAGGSGTCTGACATGGAAGTGTGTTGACAGAGAGTCT	2000
Qy	1827	CGTGGCTCTCCAGGGGAGTGTTCCTGTCTGTCTGTATTAAGAGTCAAGGAGACACCTT	1886
Db	2001	CCTGGCTCTCAGCGAGGTGTATTATGTCTGTGATTACAGAACTACTGGAATACTGT	2060
Qy	1887	TCTGTGTTACTCTGTGGCTTCCGACGCTCCGTTAGCTGACCTTATCAGTATCGGCAGAA	1946
Db	2061	GCCTGTATTCTGTCTGTCTGTCTGTATTAAGAGTCAAGCTACCGATCGTCAGAA	2120
Qy	1947	CATTCTGCTTCCAGTGTCCAGTGGCGGTAAATATCCGCTCCTGACGTCTACTGGGGCG	2006
Db	2121	CATTCTGCTTCTGAGTGGCAGTGAACATCCCGTCCGTGACTGCTACTGGGGCGG	2180
Qy	2007	TAACTGGCGACTCAGGTGAAAGCTCACCGCCATGAAATTCATATCATCTCTGAAACA	2066
Db	2181	TAACTGTGCACTCAGGTGAAGCTCACATGCAATGAATTCATTCATCATCTCTGAGCA	2240
Qy	2067	GACAAGGTTTCAAAACTAAGCATCCAGAGCCCTGAGCAGGTTTTCAGCACTGGAGGTGAA	2126
Db	2241	AACAAGGTTTCAAGAACTCAATATTGGAGAGACACGTGAACAGATTCACAGACA - CAGGA	2299
Qy	2127	GAGAGCGTGTTTTAAATAACAGAGACAAGCAGCTCAAGGTGTTTTTACAGCCCCCTGAG	2186
Db	2300	GGCAATGTTTTTTAAAGGCCCCATACAACTATTGAAGGCCATTTTACAGTTTTTCCCT	2359
Qy	2187	GGAAAGGACGCAAGGCTCTCCGACAGGTGCTCTGGGGTGAATCTTCTCTGTGAGAGCTTTTAC	2246

Db 2360 ACCGACACTGGGGGTTCTGGTCAGGCATCGGGGTGTGAATTTCCCTTTGTGGAGCGTC 2419
QY 2247 CCTCTGAGTGAGACCTCTCCACAGAGCCCCGGGGCGCGACGCCCGCTCTGTGTAGCGC 2306
Db 2420 ATCGTCAGTGACACACACACACCGTCCGCCAGTAAC-CTGGGAGGCGCTCTGGCCACACC 2478
QY 2307 TGGCAGGGGCTGTGTGGCATCAGCAGAGA-----GACGAAGCC 2347
Db 2479 AGCCTTCTAGTCGGTGTAAAGAGATCAGTAGACTGAGTTCCTCCTCCCTCCACTGGACC 2538
QY 2348 TTTCTGTAAACATCGGGCGCTCCCGCCGACAGAGGGCGAGTTTGTCTTTGTATCATTTTCC 2407
Db 2539 GCTCTCCCTCTGTGTGCTGAGGACGTGACAGTGCAGGTTTGTACTTTTGTATCTTTT-- 2596
QY 2408 GAAACTACAGTTAAAGCAGAAGTCTGTTTTCAGAAAAAGTTTTCAGGGAGAAGGGCAAGT 2467
Db 2597 ----CTACCTTGTACCTTGTACTATGTTTAAATGGGAAGCTCATAGAAAAGTTGTGAGC 2652
QY 2468 TTATCAAAAACATTTTTCAGGAG--AAGGGACATAAAGTTTACAGCCTACAGGACG-TA 2524
Db 2653 TGTGTACAGCGCATCTCCGAGAGGAAGGAAGCATACGTTTACAGCCTCCAGGGGTACA 2712
QY 2525 CACAATATCTGTGTGGGAAACACAGCATTTTATCTATTTTATTTTATTTTATTTAGGTT 2584
Db 2713 TAGGATAGTCTCCTGTCTGAAAGACACAGCCTTTTGTATATTTTATTTTATTAAGTT 2772
QY 2585 TGTGCTTATCTCTTAATAGATTTAAATGTCACAACTGTAGCACAATAATAATTT 2644
Db 2773 TGGTGCTTATCTCTTAATAGATTTAAATGTTAAATATTAATACTAGACAACTATTAATTT 2832
QY 2645 ATRATTTACAAATTCACAAAAA----- 2675
Db 2833 ATRATTTACAAATTCACAAAAATAGGAATA 2863

RESULT 3
AL539798 1201 bp mRNA linear EST 31-MAY-2003
LOCUS AL539798 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CSDF031YA23 5-PRIME, mRNA sequence.
ACCESSION AL539798
VERSION AL539798.2 GI:31264361
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1201)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polaves, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12869349.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6792.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSDF031YA23&cluster=6792.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSDF031AA12QP1.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDF031YA23"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"

/clone lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN		Query Match	34.3%; Score 919; DB 9; Length 1201;
		Best Local Similarity	96.8%; Pred. No. 1.1e-142;
		Matches	959; Conservative 9; Mismatches 19; Indels 4; Gaps 3;
QY	61	GTTCGGCGGGGGGGGGTGTGAATCCCGATGGAGCGGCCGAGGAGGACAGTCG	120
DB	67	GTTCGGCGGGGGGGGGTGTGAATCCCGATGGAGCGGCCGAGGAGGACAGTCG	126
QY	121	CGCGCCCGCAGCCCTGGGGAGCGCTCTCGCTCTGGGCGCGGAGGAGGCGGCGCAC	180
DB	127	CGCGCCCGCAGCCCTGGGGAGCGCTCTCGCTCTGGGCGCGGAGGAGGCGGCGCAC	186
QY	181	GTCTCTCTGAGGAAGCGGAGTGGACCATCGGGCGGAGAGAGGTGGACCTTCCCTTC	240
DB	187	GTCTCTCTGAGGAAGCGGAGTGGACCATCGGGCGGAGAGAGGTGGACCTTCCCTTC	246
QY	241	CCGAGCAATAAATGCTCTCTGAGATCACTGTAGAAATGTAGTGATGATCAAAATCAGGT	300
DB	247	CCGAGCAATAAATGCTCTCTGAGATCACTGTAGAAATGTAGTGATGATCAAAATCAGGT	306
QY	301	CAGGTGACACTGGAATACCAAGCACCAGTGGAAAGTGAATTAACAAGTGAAGTTGTT	360
DB	307	CAGGTGACACTGGAATACCAAGCACCAGTGGAAAGTGAATTAACAAGTGAAGTTGTT	366
QY	361	AGAAGCAGACATGCCCTTACAGCTGGGATGCTACTTCTGCTGGTGTACAGGAGAAAT	420
DB	367	AGAAGCAGACATGCCCTTACAGCTGGGATGCTACTTCTGCTGGTGTACAGGAGAAAT	426
QY	421	GAACCGGAACACAACTGGGATACCTCTATGATCTTTAAGTCAAAAGCAAGGCATGACA	480
DB	427	GAACCGGAACACAACTGGGATACCTCTATGATCTTTAAGTCAAAAGCAAGGCATGACA	486
QY	481	CAAGAAATCTTTGAAGCTCAACAGGAATATGTTTCCATGGACCAAGATACCTCAGGT	540
DB	487	CAAGAAATCTTTGAAGCTCAACAGGAATATGTTTCCATGGACCAAGATACCTCAGGT	546
QY	541	GCAGGTGCAGGGGAGGGGCGGATCCCCGGGTCCCTCGCTCGCGCCGCCACTCAGGTG	600
DB	547	GCAGGTGCAGGGGAGGGGCGGATCCCCGGGTCCCTCGCTCGCGCCGCCACTCAGGTG	606
QY	601	TGCTTTGAGGAACACAGGCCATCAACATCGATCGATCTTTCCCGACAGCTTCGGCC	660
DB	607	TGCTTTGAGGAACACAGGCCATCAACATCGATCGATCTTTCCCGACAGCTTCGGCC	666
QY	661	TCTTCCAGGAGCTTCTCTCGAGGCGAGAGCGTTCTCAGTTGTGGGTCTGGGGT	720
DB	667	TCTTCCAGGAGCTTCTCTCGAGGCGAGAGCGTTCTCAGTTGTGGGTCTGGGGT	726
QY	721	GGTGGCATCTCCCTTAAAGGAAGTGTCCCTCTGTGCGCAAGTGAATGAATCTCCAGCTTT	780
DB	727	GGTGGCATCTCCCTTAAAGGAAGTGTCCCTCTGTGCGCAAGTGAATGAATCTCCAGCTTT	786
QY	781	GCCTCAGCTCTCCAGACAAAGACTGGTCTCTTTTCTGCTTGGAAATCCCGAGATCAG	840
DB	787	GCCTCAGCTCTCCAGACAAAGACTGGTCTCTTTTCTGCTTGGAAATCCCGAGATCAG	846
QY	841	GAGGATTTGAGCCCGTGAAGAGAAATGAGAGGAGATGGGACCTTGACCTGACCGG	900
DB	847	GAGGATTTGAGCCCGTGAAGAGAAATGAGAGGAGATGGGACCTTGACCTGACCGG	906
QY	901	CAGTTGTGTGTCACAAACCGGTAGAAATGCCAAACCGTCCACAGAGACGTACAGCA	960
DB	906	CAGTTGTGTGTCACAAACCGGTAGAAATGCCAAACCGTCCACAGAGACGTACAGCA	965
QY	961	CGCGTGGAGGACAGCAAGATGGAGGAGACGCTGACATCATCTCTGCCAGGACCTG	1020

DB	966	GCCTGCTGGGAGCAAGAAAGATGGAGRG-MSCTGACATGCATM--TCTGCAAGACCTG	1022
QY	1021	CTGCAGCAGTCGCTGAGTTTGACGCCCTGCA	1051
DB	1023	CTGCACMTGCTGAGTTTSGACCCGCGATGCM	1053
RESULT 4			
AL539797/c			
LOCUS			
DEFINITION			
AL539797 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone			
EST 31-MAY-2003			
CS0DF031YA23 3-PRIME, mRNA sequence.			
ACCSSION			
AL539797			
VERSION			
AL539797.2 GI:31264360			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
REFERENCE			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 1201)			
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
On Feb 15, 2001 this sequence version replaced gi:12869347.			
COMMENT			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 6792.r For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0DF031YA12NP1&cluster=6792.r. Contact :			
Feng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ invitrogen Corporation 1600			
Faraday Avenue Genoscope sequence ID : CS0DF031YA12NP1.			
FEATURES			
Location/Qualifiers			
1..1201			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0DF031YA23"			
/tissue_type="FETAL BRAIN"			
/dev_stage="fetal"			
/clone_lib="Homo sapiens FETAL BRAIN"			
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA			
was primed with a NotI-oligo(dT) primer. Five prime end			
enriched, double-strand cDNA was digested with Not I and			
cloned into the Not I and EcoRV sites of the pCMVSPORT 6			
vector. Library was not normalized."			
ORIGIN			
Query Match			
34.0%; Score 911.8; DB 9; Length 1201;			
Best Local Similarity			
91.2%; Pred. No. 1.8e-141;			
Matches			
973; Conservative 41; Mismatches 47; Indels 6; Gaps 5;			
QY	1580	CGGAGCGGAGCAGGACCGGGTGTCCCTCTCAGCTGCGGTGTGCTGCGAGCCTT	1639
DB	1061	CGGAGCGGAGCAGGACCGGGTGTCCCTCTCAGCTGCGGTGTGCTGCGAGCCTT	1002
QY	1640	TCTGCCACCTTACTTGGGGTGTGACCGGACCGGCTCTCAGCTGCGTGGCCCCGTTT	1699
DB	1001	TCTGCCACCTTACTTGGGGTGTGACCGGACCGGCTCTCAGCTGCGTGGCCCCGTTT	945
QY	1700	GTGAGCTCAACCTGGGTGCAAGTGTCTGACCGGCTGTGACCAACAGCTACGAGT	1759
DB	944	GKRGCTCAACCTGGGTGCAAGTGTCTGACCGGCTGTGACCAACAGCTACGAGT	885
QY	1760	CAGACATCTCAAGAAATTTACCTGCGACCGAGAGTTTGACATGAAAAACATGTTGACCG	1819
DB	884	CAGACATCTCAAGAAATTTACCTGCGACCGAGAGTTTGACATGAAAAACATGTTGACCG	825

QY 1820 AGAGCCTCGTGGCTCTCCAGCGGGAGTGTTCCTGCTGTGATTAACAGAGTCACGGGAG 1879
 Db 824 AGAGCCTCGTGGCTCTCCAGCGGGAGTGTTCCTGCTGTGATTAACAGAGTCACGGGAG 765
 QY 1880 ACACCGTTCGTGTACTGCTGTGGCTGGCGAGCTTCGGTGAAGTGAAGTATCAGTATC 1939
 Db 764 ACACCGTTCGTGTACTGCTGTGGCTGGCGAGCTTCGGTGAAGTGAAGTATCAGTATC 705
 QY 1940 GGCAGAACATTCCTGCTTCGAGTTGCGAGTGGCGGCTTAACATCCGCTCTGAGTGTACT 1999
 Db 704 GGCAGAACATTCCTGCTTCGAGTTGCGAGTGGCGGCTTAACATCCGCTCTGAGTGTACT 645
 QY 2000 GGGCGCTTAATCGCGACCTCAGTGAAGCTCAGCAGCGCATGAATCAATCAATATCT 2059
 Db 644 GGGCGCTTAATCGCGACCTCAGTGAAGCTCAGCAGCGCATGAATCAATCAATATCT 585
 QY 2060 GTGAACAGACAGGTTCAAAAACCTAAGCATTCAGAGGCTTCAGAGCTTCAGAGCTGG 2119
 Db 584 GTGAACAGACAGGTTCAAAAACCTAAGCATTCAGAGGCTTCAGAGCTTCAGAGCTGG 525
 QY 2120 AGGTGAAGAGCGGTGTTTTTAAATAACAGACAGCAGCTCAAGTGTGTTTTTCAAGCC 2179
 Db 524 AGGTGAAGAGCGGTGTTTTTAAATAACAGACAGCAGCTCAAGTGTGTTTTTCAAGCC 465
 QY 2180 CCCTGAGGAGGAGCAGCGGTCTCCGACAGGTGCTCTGGGTGACTCTTCTGTGAGC 2239
 Db 464 CCCTGAGGAGGAGCAGCGGTCTCCGACAGGTGCTCTGGGTGACTCTTCTGTGAGC 405
 QY 2240 TTTTACCTCTGAGTGAGACCTCCCGACAGCGCGGGCGGCGAGCGCGCTCTCTG 2299
 Db 404 -TTTACCTCTGAGTGAGACCTCCCGACAGCGCGGGCGGCGAGCGCGCTCTCTG 347
 QY 2300 TGAGCGCTGGCGAGGCTCTGTGGTGGCATCAGCAGCAGACGAGCCTTCTGTAACTAT 2359
 Db 346 TGAGCGCTGGCGAGGCTCTGTGGTGGCATCAGCAGCAGACGAGCCTTCTGTAACTAT 287
 QY 2360 GCGCGCTCCCGCGAGGAGCGGCTTTGCTCTTTTGTATCATTTTCCGAACTACAGTT 2419
 Db 286 GCGCGCTCCCGCGAGGAGCGGCTTTGCTCTTTTGTATCATTTTCCGAACTACAGTT 227
 QY 2420 AAAGCAGAGTCTGTTTTTCAAGAAAGTTTCAAGGAGAGGCGCAAGTTTATCAAAAACA 2479
 Db 226 AAAGCAGAGTCTGTTTTTCAAGAAAGTTTCAAGGAGAGGCGCAAGTTTATCAAAAACA 167
 QY 2480 TTGTTTCAGGAGAGGAGCATAAGTTTACAGCTACAGGACGTACACATATCTCTGCTG 2539
 Db 166 TTGTTTCAGGAGAGGAGCATAAGTTTACAGCTACAGGACGTACACATATCTCTGCTG 107
 QY 2540 CTGGAAACCCAGCAGCATTTTATCTATTTTATTTTATTAAGTTTGGTCTTATCTTCT 2599
 Db 106 CTGGAAAGSAGCGGGTTTTATSTDTTTTATTTTATTAAGTTTGGTCTTATCTTCT 47
 QY 2600 AATAAGATTTAAATGTCAAACTGTAGCACAATAATAATTTAT 2646
 Db 46 AATAAGATTTAAATGTCAAACTGTAGCACAATAATAATTTAT 1

RESULT 5
 BQ876252
 LOCUS BQ876252 868166 NIH_MGC_43 Homo sapiens cdna clone IMAGE:6376937
 DEFINITION AGNCOURT 868166 NIH_MGC_43 Homo sapiens cdna clone IMAGE:6376937
 5', mRNA sequence.
 ACCESSION BQ876252
 VERSION BQ876252.1 GI:22268258
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 876)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Straubeberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ARCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LUCM2559 row: o column: 18
 High quality sequence stop: 678.
 Location/Qualifiers
 1. 876
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6376937"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_43"
 /note="Organ: eye; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 30.7%; Score 822.2; DB 13; Length 876;
 Best Local Similarity 98.3%; Fred. No. 1.4e-126;
 Matches 863; Conservative 0; Mismatches 9; Indels 6; Gaps 3;
 QY 1165 GTGGAAGCATACCTCATCCAGCATCCAGACAGAGTCCAGTGAAGAAGATGTGCAAAAGT 1224
 Db 1 GTGGAAGCATACCTCATCCAGCATCCAGACAGAGTCCAGTGAAGAAGATGTGCAAAAGT 60
 QY 1225 ATGGATGCGAGGAGGAGTTCACTCAAGACATGCTGAGCCCAAGTCAAGCGGCTTTT 1284
 Db 61 ATGGATGCGAGGAGGAGTTCACTCAAGACATGCTGAGCCCAAGTCAAGCGGCTTTT 120
 QY 1285 TCTGATGAAGAGGAGGAGTTCAAGAGACCTGTGGAGCTGTGAGCTTCAAGCTTCAAGTCAAGTCC 1344
 Db 121 TCTGATGAAGAGGAGGAGTTCAAGAGACCTGTGGAGCTGTGAGCTGTGAGCTGTGAGCTGAGTCC 180
 QY 1345 TCAGACATTAGCAGCCCATACGTCTGTGTCGCGGAGTCTCTGTAGTACAGAAAGGAGCGG 1404
 Db 181 TCAGACATTAGCAGCCCATACGTCTGTGTCGCGGAGTCTCTGTAGTACAGAAAGGAGCGG 240
 QY 1405 GCGCAGCTCCCACTGCGCCAGCACCCGAGGCGGAGCCAGAGCCCAAGCCCTGGGG 1464
 Db 241 GCGCAGCTCCCACTGCGCCAGCACCCGAGGCGGAGCCAGAGCCCAAGCCCTGGGG 300
 QY 1465 GATGACACCTCCAGCTCCGTCAGCCTGAGCAGCAGTCCAGGATTAAGTGTGCTCTG 1524
 Db 301 GATGACACCTCCAGCTCCGTCAGCCTGAGCAGTCCAGGATTAAGTGTGCTCTG 357
 QY 1525 CAAGAAGCCAGCCCTGTGACCTGTCTTCCAGCCCATGCCAGCCGAGAGCGGAG 1584
 Db 358 CAAGAAGCCAGCCCTGTGACCTGTCTTCCAGCCCATGCCAGCCGAGAGCGGAG 417
 QY 1585 CGCAGCAGGAGACCCGCGTGTGCGCCCTCAGCAGTGTGCGGTGTGCTGCGAGCTTTTTCG 1644
 Db 418 CGCAGCAGGAGACCCGCGTGTGCGCCCTCAGCAGTGTGCGGTGTGCTGCGAGCTTTTTCG 477
 QY 1645 CACCTGTACTGGGGGTGACCCGAGACCGGCTGCTACGCTGCTGCGCCCTTTTGTGAG 1704
 Db 478 CACCTGTACTGGGGGTGACCCGAGACCGGCTGCTACGCTGCTGCGCCCTTTTGTGAG 537
 QY 1705 CTCACACTGGGTGACAAAGTGTCTGACCGCGCTGCTGAACAAACAAAGTGTACAGTACAGT 1764
 Db 538 CTCACACTGGGTGACAAAGTGTCTGACCGCGCTGCTGAACAAACAAAGTGTACAGTACAGT 597

QY 1765 ATCTGAAGAAATACCTGGCAACACAGAGTTTGACATGGAAGAAACATGTTGACCGAGC 1824
Db 598 ATCTGAAGAAATACCTGGCAACACAGAGTTTGACATGGAAGAAACATGTTGACCGAGC 657
QY 1825 CTCGTGGCTCCAGGGGGAGTGTCTGCTGTCTGATACAGAGTCAGGG--AGACA 1882
Db 658 CTCATGGCTCTCCAGGGGGAGTGTCTGCTGTCTGATACAGAGTCAGGGGGAGACAC 717
QY 1883 CGGTTCTGTGTACTGTGTGGCTGGCCAGCTTCCGTGAGCTGACCTATCAGTATCGGC 1942
Db 718 CGGTTCTGTGTACTGTGTGGCTGGCCAGCTTCCGTGAGCTGACCTATCAGTATCGGC 777
QY 1943 AGAACAATCTCTCTCCAGTGGCCAGTGGCCGTAACATCCGTCCTGACTGTACTGGG 2002
Db 778 AGAACAATCTCTCTCCAGTGGCCAGTGGCCGTAACATCCGTCCTGACTGTACTGGG 837
QY 2003 -GCCGTAACTGCCGCACTCAGGTGAAGCTCACCACGC 2039
Db 838 NGGCGTAATGCCGCACTCCGTAAGAGCTCACCACGC 875

RESULT 6
BUI76559 921 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT_7940317 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143609
DEFINITION 5', mRNA sequence.
ACCESSION BUI76559
VERSION BUI76559.1 GI:22690543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 921)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL3467 row: e column: 18
High quality sequence stop: 897.
Location/Qualifiers
1..921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6143609"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-Sport6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

FEATURES
source
Location/Qualifiers
1..921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6143609"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-Sport6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 28.9%; Score 774.8; DB 13; Length 921;
Best Local Similarity 99.7%; Fred. No. 1e-118;
Matches 776; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1889 TGTGTACTGTGTGGCTGGCCAGTTCGGTGAAGCTGACCTATCAGTATCGGCAAGCA 1948
Db 1 TGTGTACTGTGTGGCTGGCCAGTTCGGTGAAGCTGACCTATCAGTATCGGCAAGCA 60
QY 1949 TTCCTGCTCCGAGTTCGAGTGGCCGTAACATCCCGTCTGCTGACTGTCTGAGGCGCGTA 2008

Db 61 TTCTGCTCCGAGTTGCCAGTGGCCGTAACATCCCGTCTGCTGACTGTCTGAGGCGCGTA 120
QY 2009 ACTGCCGCACTCAGGTGAAGCTCACCACGCCATGAATTCATATCATATCTGTGAACAGA 2068
Db 121 ACTGCCGCACTCAGGTGAAGCTCACCACGCCATGAATTCATATCATATCTGTGAACAGA 180
QY 2069 CAAGTTCAAAAACTAAGCATCCAGAGCCCTGAGCAGCTTTTTCAGCACTGGAGGTGAAGA 2128
Db 181 CAAGTTCAAAAACTAAGCATCCAGAGCCCTGAGCAGCTTTTTCAGCACTGGAGGTGAAGA 240
QY 2129 GAGCGTGTGTAAAAATACAGAGCAAGCAGTCAAGGTGTGTTCACAGCCCTCGAGGG 2188
Db 241 GAGCGTGTGTAAAAATACAGAGCAAGCAGTCAAGGTGTGTTCACAGCCCTCGAGGG 300
QY 2189 AAGGAGCCAGGGTCTCGGACAGGTCTCTGGGGTGAATCTCTCTGTGGAGCTTTTACCC 2248
Db 301 AAGGAGCCAGGGTCTCGGACAGGTCTCTGGGGTGAATCTCTCTGTGGAGCTTTTACCC 360
QY 2249 TCTGAGTGAGACCCCTCCCGAGAGCCCGGGGGCCGAGCCCGCCCTCTGCTGAGGCGTG 2308
Db 361 TCTGAGTGAGACCCCTCCCGAGAGCCCGGGGGCCGAGCCCGCCCTCTGCTGAGGCGTG 420
QY 2309 GCGAGGGTCTGCTGTGGCATCAGCAGAGAGAGAGCCCTTCTGTAAACATGCGGCGCTC 2368
Db 421 GCGAGGGTCTGCTGTGGCATCAGCAGAGAGAGAGCCCTTCTGTAAACATGCGGCGCTC 480
QY 2369 CCGCCGAGAGGGGCGAGTTTCTCTTTGTACATTTTCCGAAACTACAGTTAAAGCAGAA 2428
Db 481 CCGCCGAGAGGGGCGAGTTTCTCTTTGTACATTTCCGAAACTACAGTTAAAGCAGAA 540
QY 2429 GTCTGTTTTCAGGAAAGTTTCAAGGAGAGGGGAGGTTTATCAAAACATGTTTTCAG 2488
Db 541 GTCTGTTTTCAGGAAAGTTTCAAGGAGAGGGGAGGTTTATCAAAACATGTTTTCAG 600
QY 2489 GAGAGGGGAGCATTAAGTTTACAGCTACAGAGCGTACACAAATATCTGCTGCGGAAA 2548
Db 601 GAGAGGGGAGCATTAAGTTTACAGCTACAGAGCGTACACAAATATCTGCTGCGGAAA 660
QY 2549 CCAGCAGATTTATCTATTTTATTTTAAAGTTTGGTGTCTATCTTCTAATAGATT 2608
Db 661 CCAGCAGATTTATCTATTTTATTTTAAAGTTTGGTGTCTATCTTCTAATAGATT 720
QY 2609 TAAATGTCAACAACCTGAGCAAAATAATAATTTATTAATTTTACAAATTTGACAAAA 2665
Db 721 TAAATGTCAACAACCTGAGCAAAATAATAATTTATTAATTTTACAAATTTGACTAAA 778

RESULT 7
BQ215731 870 bp mRNA linear EST 02-MAY-2002
LOCUS BQ215731
DEFINITION AGENCOURT_7549119 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059147
5', mRNA sequence.
ACCESSION BQ215731
VERSION BQ215731.1 GI:20397131
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 870)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1AM1325 row: f column: 12
High quality sequence stop: 601.
Location/Qualifiers
1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6059147"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 68"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 28.6%; Score 765.4; DB 13; Length 870;
Best Local Similarity 98.8%; Pred. No. 3.9e-117;
Matches 793; Conservative 0; Mismatches 6; Indels 4; Gaps 2;

QY 916 CAACCGGTAGAAATGCCAAACCGTCCACGAGACGTCAGACAGCGGCTGGGAGGCCA 975
DB 1 CAACCGGTAGAAATGCCAAACCGTCCACGAGAGCGTCAGACAGCGGCTGGGAGGCCA 60

QY 976 GACAAAGTAGGAGAGAGCGTGCATGTCATCTGCCAGGACCTGTCGACGACTGGTG 1035
DB 61 GACAAAGTAGGAGAGAGCGTGCATGTCATCTGCCAGGACCTGTCGACGACTGGTG 120

QY 1036 AGTTTGAGCGCCCTGCATGCACACGTTCTGGCGGCTTGCTACTCGGCTGATGAGAGCGC 1095
DB 121 AGTTTGAGCGCCCTGCATGCACACGTTCTGGCGGCTTGCTACTCGGCTGATGAGAGCGC 180

QY 1096 TCGTCCCTGTGTCTACTCGCTGCTGCGTGAGCGGATCTGTAAACCAACATCTCTC 1155
DB 181 TCGTCCCTGTGTCTACTCGCTGCTGCGTGAGCGGATCTGTAAACCAACATCTCTC 240

QY 1156 AACAACCTCTGGAAGCATACCTCATCCAGCATCCAGCAAGAGTGCAGTGAAGAGAT 1215
DB 241 AACAACCTCTGGAAGCATACCTCATCCAGCATCCAGCAAGAGTGCAGTGAAGAGAT 300

QY 1216 GTGCAAAAGTAGGATGCCAGGAATAAATCACTCAAGACATGTGCGAGCCCAAGTCAGG 1275
DB 301 GTGCAAAAGTAGGATGCCAGGAATAAATCACTCAAGACATGTGCGAGCCCAAGTCAGG 360

QY 1276 CGGTCTTTTCTGATGAGAGGAGTTTCAGAGGACCTGCTGGAGCTGTGACAGCTTGAC 1335
DB 361 CGGTCTTTTCTGATGAGAGGAGTTTCAGAGGACCTGCTGGAGCTGTGACAGCTTGAC 420

QY 1336 AGTGAGTCTCAGACATTAGCCAGCCATACGTCTGTGCCGCGAGTGTCTGAGTACAGA 1395
DB 421 AGTGAGTCTCAGACATTAGCCAGCCATACGTCTGTGCCGCGAGTGTCTGAGTACAGA 480

QY 1396 AGGAGGCGGCGAGCTCCCATCGCCAGCACCCGAGGCGAGCGAGGCCACAG 1455
DB 481 AGGAGGCGGCGAGCTCCCATCGCCAGCACCCGAGGCGAGCGAGGCCACAG 540

QY 1456 GCCCTGGGGATGACACCTCCAGCTCCGTGACGCTGACGACAGCAGTTCACAGATTACGTG 1515
DB 541 GCCCTGGGGATGACACCTCCAGCTCCGTGACGCTGACGACGCTGACGACGCTGACGCTG 597

QY 1516 TGCCCTCTGCAAGGAGACGACGCTGTGCACTGCTGCTTCCAGCCCATGCCGACCGG 1575
DB 598 TGCCCTCTGCAAGGAGACGACGCTGTGCACTGCTGCTTCCAGCCCATGCCGACCGG 657

QY 1576 AGAGCGGAGCGGAGGAGGACCCGCGTGTGCCCGCTCAGCAGTGTGCGGTCTGCTGTCAG 1635
DB 658 AGAGTGGAGCGGAGGAGGACCCGCGTGTGCCCGCTCAGCAGTGTGCGGTCTGCTGTCAG 717

QY 1636 CCTTTTGCCACTGTACTGGGGTGCACCCGGAGCCGGCTGTCTACGCGCTGCTGG-CCCC 1694
DB 718 CCTTTTGCCACTGTACTGGGGTGCACCCGGAGCCGGCTGTCTACGCGCTGCTGGNCCCC 777

QY 1695 GTTTGTGAGCTCAACCTGGGTG 1717
DB 778 GTTTGTGAGCTCAACCTGGGG 800

RESULT 8
EQ073354
LOCUS
DEFINITION
BO073354 992 bp mRNA linear EST 02-APR-2002
AGENCOURT_7052199 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:5806143
5', mRNA sequence.
BO073354
ACCESSION
BO073354.1 GI:19902400
VERSION
BO073354
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 992)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW2046 row: h column: 16
High quality sequence stop: 670.

FEATURES
Location/Qualifiers
1. 992
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5806143"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match 28.3%; Score 758; DB 13; Length 992;
Best Local Similarity 96.6%; Pred. No. 6.3e-116;
Matches 817; Conservative 0; Mismatches 25; Indels 4; Gaps 4;

QY 551 GCGGAGGGGCGGATCCCGGGTCCCTCGTGTGCGCGCCACTCAGGTGCTTTGAGG 610
DB 1 GCGGAGGGGCGGATCCCGGGTCCCTCGTGTGCGCGCCACTCAGGTGCTTTGAGG 60

QY 611 AACCAACAGCATTCAACATCGACGTCAGACCTCTTCCCAACAGCCTCGGCTTCCACGG 670
DB 61 AACCAACAGCATTCAACATCGACGTCAGACCTCTTCCCAACAGCCTCGGCTTCCACGG 120

QY 671 AGCTTCTCTCAGGCGGAGAGCGTTCTCCAGTGTGGTCTGGGGTGGTGGCATCT 730
DB 121 AGCTTCTCTCAGGCGGAGAGCGTTCTCCAGTGTGGTCTGGGGTGGTGGCATCT 180

QY 731 CCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTGAAGTCTCCAGCTTTGCTCAGCTC 790
DB 181 CCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTGAAGTCTCCAGCTTTGCTCAGCTC 240

QY 791 TCCAGACAGAAAGACTGGTCTCTTTTGTCTGCTTGAACCCCGAGATCAGGAGGATTTGG 850

241 TCCAGACAGAAAGAGCTCGTCTCTTTTGGTGGTGAACCCAGGATCAGGAGATTGG 300
851 AGCCCGTGAAGAAAGAAATGAGAGAGATGGGACCTTGACTGAAAGGGACATGGTGG 910
301 AGCCCGTGAAGAAAGAAATGAGAGAGATGGGACCTTGACTGAAAGGGACATGGTGG 360
911 TGGCAACACCGGTGAATGAAATGCCAAACCGTCCACGAGGAGCTCAGAGCAGCGCTGGGA 970
361 TGGCAACACCGGTGAATGAAATGCCAAACCGTCCACGAGGAGCTCAGAGCAGCGCTGGGA 420
971 AGCCAGACAAAGATGAGAGAGAGCTGACATGATCATCTGCGAGGAGCTCTGCGAGACT 1030
421 AGCCAGACAAAGATGAGAGAGAGCTGACATGATCATCTGCGAGGAGCTCTGCGAGACT 480
1031 GCGTGAGTTGAGCCCTGCAATGCAACAGTTCTGCGGGGCTTGCTACTCGGGTGGATGG 1090
481 GCGTGAGTTGAGCCCTGCAATGCAACAGTTCTGCGGGGCTTGCTACTCGGGTGGATGG 540
1091 AGCGTCTGCTCCCTGTGTCTTACCTGCTCCCTGCTCCGCTGGAGCGGATCTGTAAGAACCA 1150
541 AGCGTCTGCTCCCTGTGTCTTACCTGCTCCCTGCTCCGCTGGAGCGGATCTGTAAGAACCA 600
1151 TCCTCAACAACTCTGTGAAGATACCTCATCTCCAGCATCCAGACAGAGTGGAGTGAAG 1210
601 TCCTCAACAACTCTGTGAAGATACCTCATCTCCAGCATCCAGACAGAGTGGAGTGAAG 660
1211 AAGATGTGCAAGATGATGATGCCAGGATAAAATCACTCAAG-ACATGCTGCAGGCCAAA 1269
661 AAGATGTGCAAGATGATGATGCCAGGATAAAATCACTCAAG-ACATGCTGCAGGCCAAA 720
1270 GTGAGCGGTCTTTTCTGATGAAGAGAGGAGTTTTCAGAGACCTGCT-GGAGTGTGAG- 1327
721 GTGAGCGGTCTTTTCTGATGAAGAGAGGAGTTTTCAGAGACCTGCTGGGACCTGTGAGA 780
1328 AGTTGACAGTGAG-TCTCTCAGACATTAGCCAGGACATACGCTGCTGCGGCGAGTGTCT 1386
781 AGTTGACAGTGAGTTCTCTCAACATTAGCCAGGACCTTCTTCTGGGCGCGGAGT 840
1387 GAGTAC 1392
841 GGGTCC 846

CD243458 877 bp mRNA linear EST 22-MAY-2003
AGENCY: NIH MGC 180 Homo sapiens cDNA clone
IMAGE: 30383176 5', mRNA sequence.
CD243458
CD243458.1 GI:31003922
EST.
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rcs@nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM446 row: j column: 17
High quality sequence stop: 660.
Location/Qualifiers
1..877
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30383176"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH MGC 180"
/note="Organ: Testis; Vector: pCMV-Sport6.1; Site 1: NotI;
Site 2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (invitrogen). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 28.3%; Score 757.6; DB 14; Length 877;
Best Local Similarity 96.9%; Pred. No. 7.6e-116;
Matches 772; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1527 AGGAGGACCGCCCTGTCACCTGCTTCCAGCCCATGCCCCGAGCGGAGCGGAGCG 1586
DB 4 AGGAGGACCGCCCTGTCACCTGCTTCCAGCCCATGCCCCGAGCGGAGCGGAGCG 63
QY 1587 CGAGCAGGACCGCGCTGTCGCCCTCAGCAGTGTGCGGTCTGCTCAGCCTTTTGCCA 1646
DB 64 CGAGCAGGACCGCGCTGTCGCCCTCAGCAGTGTGCGGTCTGCTCAGCCTTTTGCCA 123
QY 1647 CCTGTACTGGGGTGCACCGCGCTGCTAGCGCTGCTGCGCTGCTGCGCTGCTGCGCT 1706
DB 124 CCTGTACTGGGGTGCACCGCGCTGCTAGCGCTGCTGCGCTGCTGCGCTGCTGCGCT 183
QY 1707 CAACCTGGGTGACAAAGTGTCTGACCGCGTGTGAAACAACAGCTACAGTCAAGT 1766
DB 184 CAACCTGGGTGACAAAGTGTCTGACCGCGTGTGAAACAACAGCTACAGTCAAGT 243
QY 1767 CCTGAGCAATTTACCTGTCACCGAGGTTTTCATGCAATGCAATGCAATGCAATG 1826
DB 244 CCTGAGCAATTTACCTGTCACCGAGGTTTTCATGCAATGCAATGCAATGCAATG 303
QY 1827 CGTGGCTCTCCAGCGGGAGTGTCTGCTGTCTGATTACAGAGTCAAGGAGACACCT 1886
DB 304 CGTGGCTCTCCAGCGGGAGTGTCTGCTGTCTGATTACAGAGTCAAGGAGACACCT 363
QY 1887 TCTGTGTTACTGCTGTGGCTGCGAGCTTCCGAGAGTCAAGTCAAGTCAAGTCAAGT 1946
DB 364 TCTGTGTTACTGCTGTGGCTGCGAGCTTCCGAGAGTCAAGTCAAGTCAAGTCAAGT 423
QY 1947 CATTCCTGCTTCCAGGTTGCCAGTGGCCCTTAACATCCCTGCTGCTGCTGCTGCTG 2006
DB 424 CATTCCTGCTTCCAGGTTGCCAGTGGCCCTTAACATCCCTGCTGCTGCTGCTGCTG 483
QY 2007 TAACTGCGGCACCTCAGTGTGAAGCTCACCACCGCATGAATTCATCATATCTGTGA 2066
DB 484 TAACTGCGGCACCTCAGTGTGAAGCTCACCACCGCATGAATTCATCATATCTGTGA 543
QY 2067 GACAAAGTTCAAAAATAAGCATCCAGAGCGCTGAGCAGCTTTTTCAGCACTGGAGT 2126
DB 544 GACAAAGTTCAAAAATAAGCATCCAGAGCGCTGAGCAGCTTTTTCAGCACTGGAGT 603
QY 2127 GAGAGCGTGTGTTTAAATAACAGAGACAGCAGTCAAGGTGTTTTCAGCAGCCCTG 2186
DB 504 GAGAGCGTGTGTTTAAATAACAGAGACAGCAGTCAAGGTGTTTTCAGCAGCCCTG 663
QY 2187 GGAAGGAGACCGAGGCTCTCCGACAGGTGTCTGCGGTGAGCTCTTCTGTGGAGCTTTT 2246
DB 664 GGAAGGAGACCGAGGCTCTCCGACAGGTGTCTGCGGTGAGCTCTTCTGTGGAGCTTTT 723
QY 2247 CCTCTGAGTGAGACCTCTCCCGAGAGCGCGGGGGCGGAGCGCCCTCTCTGTGAGCG 2306
DB 724 CCTCTGAGTGAGACCTCTCCCGAGAGCGCGGGGGCGGAGCGCCCTCTCTGTGAG 783
QY 2307 TGGGCGAGGCTCTGTTGT 2323
DB 784 GCGCTTGGGCGAGGCT 800


```
RESULT 10
BUI148471
LOCUS
DEFINITION AGENCOURT 8675647 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380102
5' RNA sequence.
ACCESSION BUI148471.1 GI:22662003
VERSION BUI148471.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2568 row: c column: 15
High quality sequence stop: 620.
FEATURES
source
Location/Qualifiers
1..910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380102"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 27.68; Score 740.6; DB 13; Length 910;
Best Local Similarity 91.18; Pred. No. 5e-113;
Matches 858; Conservative 0; Mismatches 40; Indels 44; Gaps 5;
Qy 12 ACAGAGCGCGCAATCTCTTTGACAGCGCGCGCGCGCGCGGTTCCGGTTCCGGCGCG 71
Db 5 AGGAGCGCGCAATCTCTTTGACAGCGCGCGCGCGCGCGGTTCCGGTTCCGGCGCG 64
Qy 72 GCGCGGATGTGAATCCCGATGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 131
Db 65 GCGCGGATGTGAATCCCGATGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 124
Qy 132 GCGCGGAGCGGCTCTCTGCGTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 191
Db 125 GCGCGGAGCGGCTCTCTGCGTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184
Qy 192 GAACGCGGAGTGGACCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 251
Db 185 GAACGCGGAGTGGACCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 244
Qy 252 ACTGCTCTCGAGATCACTGTAGAAATTTAGTGGATGAAATTCAGTTCAGTGACACT 311
Db 245 ACTGCTCTCGAGATCACTGTAGAAATTTAGTGGATGAAATTCAGTTCAGTGACACT 304
Qy 312 GGAAGATACCAAGCAGTGGAAAGATTAACAAGCTGAAGTTCTTAAGAGCAGAC 371
```

```
Db 305 GGAAGATACCAAGCAGTGGAAAGATTAACAAGCTGAAGTTCTTAAGAGCAGAC 364
Qy 372 ATGCCCTTTACAGACTGGGGATGTCACTACTTGGTGTACAGGAAGATGACCGGACA 431
Db 365 ATGCCCTTTACAGACTGGGGATGTCACTACTTGGTGTACAGGAAGATGACCGGACA 424
Qy 432 CAACGTGGCATACCTCTATGAATCTTTAAAGTGAAGCAAGGCATGACACAAGATCCTT 491
Db 425 CAACGTGGCATACCTCTATGAATCTTTAAAGTGAAGCAAGGCATGACACAAGATCCTT 484
Qy 492 TGAAGCTAACAGGAATAATGTGTTCCATGGACCAAGATACCTCAGGTGAGGTGAGG 551
Db 485 T-----GATACCTCAGGTGAGGTGAGG 508
Qy 552 GCGAGGCGCGATCCCGGGTCCCTCCGTCCCGCCGCGCCACTCAGGTGTCTTTGAGGA 611
Db 509 GCGAGGCGCGATCCCGGGTCCCTCCGTCCCGCCGCGCCACTCAGGTGTCTTTGAGGA 568
Qy 612 ACCCAGGCATCAACATCGACGTGAGACCTCTTCCCAACAGCTCGGCTCTTCCACGGA 671
Db 569 ACCCAGGCATCAACATCGACGTGAGACCTCTTCCCAACAGCTCGGCTCTTCCACGGA 628
Qy 672 GCCTTCTCTCGAGGCGAGAGCGTTCTCAGTTGTGGTCTGGGGTGGTGGCATCTC 731
Db 629 GCCTTCTCTCGAGGCGAGAGCGTTCTCAGTTGTGGTCTGGGGTGGTGGCATCTC 688
Qy 732 CCTAAAGCAAGTGTCTCTGTGGCAAGTGATGAAGTCTCCAGCTTTGCTCAGCTCT 791
Db 689 CCTAAAGCAAGTGTCTCTGTGGCAAGTGATGAAGTCTCCAGCTTTGCTCAGCTCT 748
Qy 792 CCCACAGAGAAGACT-GCGTCTCTTTTCGTGTGGAAACCCAGGATCAGAGGATTTGG 850
Db 749 CCCAACAGAGAAGACTGGGGTCTTTTCGTGTGGAAACCCAGGATCAGAGGATTTG 808
Qy 851 AGCCCGTGAAGAGAAATGA----GAGGAGATGGGACCTTG--ACCTGAACGGGCGAGT 904
Db 809 GGAGCGCGTGAAGAGAAATGACAGAGATGGGACCTTGACCTGAACGGGCGAGT 868
Qy 905 TGTGTTGTCGACCAAC-CCGTAGAAATCCCAACCGCTCCAC 945
Db 869 GTGTGTTCCCAACACCGGGTAGAAATGCCCAACCGGGCCC 910
RESULT 11
BUI148471
LOCUS
DEFINITION 602458365F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4580633 5',
mRNA sequence.
ACCESSION BG395714
VERSION BG395714.1 GI:13289162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 840)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1300 row: a column: 18
High quality sequence stop: 775.
FEATURES
source
Location/Qualifiers
1..840
/organism="Homo sapiens"
/mol_type="mRNA"
```

ORIGIN		/db_xref="taxon:9606" /clone="IMAGE:4580833" /tissue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_16" /notes="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."									
Query Match	27.2%	Score 727.4;	DB 12;	Length 840;							
Best Local Similarity	95.8%;	Pred. No. 8e-111;									
Matches 779;	Conservative 0;	Mismatches 31;	Indels 3;	Gaps 3;							
QY	1294	GAAGGAGTTTCAGAGACCTGCTGGAGCTGTCTCAGACGTTGACAGTACAGTCTCTCAGACATT	1353								
DB	2	GAAGGAGTTTCAGAGACCTGCTGGAGCTGTCTCAGACGTTGACAGTACAGTCTCTCAGACATT	61								
QY	1354	AGCCAGCATPACGTCTGTCTGCCGCGAGTCTCTGAGTACAGAAAGGACGCGCGCGACGCT	1413								
DB	62	AGCCAGCATPACGTCTGTCTGCCGCGAGTCTCTGAGTACAGAAAGGACGCGCGCGACGCT	121								
QY	1414	CCCCACTGCCACACCCCGAGGCGAGCCAGGAGCCCCACAGGCCCTTGGGGGATGCACCC	1473								
DB	122	CCCCACTGCCACACCCCGAGGCGAGCCAGGAGCCCCACAGGCCCTTGGGGGATGCACCC	181								
QY	1474	TCCACGTCCTGTACGCTGACGACAGCAGTCCAGGATTCAGTGTCCTCTGCAAGGAAGC	1533								
DB	182	TCCACGTCCTGTACGCTGACGACAGCAGTCCAGGATTCAGTGTCCTCTGCAAGGAAGC	241								
QY	1534	CAGCCCTGTGCACCTGCTCTTCCAGCCCATGCCGACCGGAGAGCGGAGCGGAGCAG	1593								
DB	242	CAGCCCTGTGCACCTGCTCTTCCAGCCCATGCCGACCGGAGAGCGGAGCGGAGCAG	301								
QY	1594	GACC CGGTGTGCGCCCTCAGCAGTGTGCGGTCTGCTCGAGCCTTTCTGCCACCTGTAC	1653								
DB	302	GACC CGGTGTGCGCCCTCAGCAGTGTGCGGTCTGCTCGAGCCTTTCTGCCACCTGTAC	361								
QY	1654	TGGGGTGTGACCCCGGACCGGTGTGTCAGCGTGTGCGTGTGCGTGTGAGCTCAACCTG	1713								
DB	362	TGGGGTGTGACCCCGGACCGGTGTGTCAGCGTGTGCGTGTGCGTGTGAGCTCAACCTG	421								
QY	1714	GGTGACAAGTGTCTGGACGGGTGCTCAACAAACAACAGTACGAGTCAGACATCTCTGAAG	1773								
DB	422	GGTGACAAGTGTCTGGACGGGTGCTCAACAAACAACAGTACGAGTCAGACATCTCTGAAG	481								
QY	1774	AATTACCTGGCAACACAGAGTTTGACATGGAAAAACATGTTGACCGAGAGCCTCGTGGCT	1833								
DB	482	AATTACCTGGCAACACAGAGTTTGACATGGAAAAACATGTTGACCGAGAGCCTCGTGGCT	541								
QY	1834	CTCCAGCGGGAGTGTCTCTGTCTGTGATTTACAGAGTCAGCGGAGACACCGTTCTGTGT	1893								
DB	542	CTCCAGCGGGAGTGTCTCTGTCTGTGATTTACAGAGTCAGCGGAGACACCGTTCTGTGT	601								
QY	1894	TACTGCTGTGGCCTGCGCAGCTTCCGTGAGCTGACCTATCAGTATCGCGAGAACATTCTCT	1953								
DB	602	TACTGCTGTGGCCTGCGCAGCTTCCGTGAGCTGACCTATCAGTATCGCGAGAACATTCTCT	661								
QY	1954	GCTTCCGAGTTGCCAGTGGCCGTTAACATCCGTCCTGACTCTACTGGGCGCGTAACCTGC	2013								
DB	662	GCTTCCGAGTTGCCAGTGGGCGTTAACATCCGTCCTGACTCTACTGGGCGCGTAACCTGC	721								
QY	2014	CGCACTCA-GGTGAAGCTC-ACCACGCCATGAATTCATATCATATCTGT-GAACAGACA	2070								
DB	722	GCACCTCAGGTGAAGCTCAACAGGCCATGAATTCATATCATATCTGTGACACGAGAC	781								
QY	2071	AGGTTCAAAAACTAAGCATTCAGAGGCCCTGAG	2103								

```

FEATURES
source
Location/Qualifiers
1..953
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:6371968"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

```

ORIGIN	Query Match	27.1%;	Score 726.8;	DB 13;	Length 953;
	Best Local Similarity	93.0%;	Prod No. 9.6e-111;		
	Matches 842;	Conservative 0;	Mismatches 18;	Indels 45;	Gaps 16;
Qy	45	CGCGAGCCGGTTCGGGTTCCGGGTCGGGCGGGGGATGTGAATCCGATGAGCGGCCCGA	104		
Db	1	CGCGAGCCGGTTCGGGTTCCGGGTCGGGCGGGGGATGTGAATCCGATGAGCGGCCCGA	60		
Qy	105	GGAGGGCAAGCAGTCGCGCGCGCGGACGCTCTGGGACGGCTCTCGTGGTCTGGGCGCGGA	164		
Db	61	GGAGGGCAAGCAGTCGCGCGCGCGGACGCTCTGGGACGGCTCTCGTGGTCTGGGCGCGGA	120		
Qy	165	GGAGGGCGAGCCGCACTCTCTTGAGGAAGCGGGAGTGGACCAATCGGCGGAGACGAGG	224		
Db	121	GGAGGGCGAGCCGCACTCTCTTGAGGAAGCGGGAGTGGACCAATCGGCGGAGACGAGG	180		
Qy	225	TTGGGACCTTCTCTTCCCGCAGCAATAAATGGTCTCTGGAGATCACTGTGTAGAATTGTAGT	284		
Db	181	TTGGGACCTTCTCTTCCCGCAGCAATAAATGGTCTCTGGAGATCACTGTGTAGAATTGTAGT	240		
Qy	285	GGATGAAAAATCAGGTCAAGTGGACATCTGGAAAGATACCAGCACCAAGTGGACAGTATTA	344		
Db	241	GGATGAAAAATCAGGTCAAGTGGACATCTGGAAAGATACCAGCACCAAGTGGACAGTATTA	300		

345 CRAGCTGAGGTGTGTTAAGAACGACGACATGCCCTTTACAGACTGGGGATGTCATCTACTT 404
Db |||||
301 CRAGCTGAGGTGTGTTAAGAACGACGACATGCCCTTTACAGACTGGGGATGTCATCTACTT 360
Qy GTGTGACAGGAAGATGAACCGGAACACAACTGGGCATACCTCTATGATCTTTAAGTGA 464
Db GTGTGACAGGAAGATGAACCGGAACACAACTGGGCATACCTCTATGATCTTTAAGTGA 420
Qy AAAGCAGGACATGACACAGATCTTTGAAGCTAACAGGAATGTTCCATGGGAC 524
Db AAAGCAGGACATGACACAGATCTTTGAAGCTAACAGGAATCTTTT 448
Qy CAAAGATACCTCAGGTGACAGGTGACAGGGGAGGGGCGGATCCCGGTCCTCCGTC 584
Db ---GATACCTCAGGTGACAGGTGACAGGGGAGGGGCGGATCCCGGTCCTCCGTC 504
Qy GCCCGCATCAGGTGTGTTGAGGAACACAGCCATCAACATGACATGACATCTT 644
Db GCCCGCATCAGGTGTGTTGAGGAACACAGCCATCAACATGACATGACATCTT 564
Qy CCCACAGCTCGGCTCTTCCAGGAGCTTCTCTGAGGAGCTTCTCTGAGGAGAGCTTCCAG 704
Db CCCACAGCTCGGCTCTTCCAGGAGCTTCTCTGAGGAGAGCTTCTCTGAGGAGAGCTTCCAG 624
Qy TTGTGGGTCT- GGGGTGTGTGATCTCCCTTAAAGGAAGTGGTCCCTCTGTCGTCG 763
Db TTGTGGGTCTGGGGGTGGTGGCATCTCCCTTAAAGGAAGTGGTCCCTCTGTCGTCG 684
Qy ATGAGTCTCCAGCTTGGCTCAGCTCTCCAGAGAGAGTGGTCCCTCTGTCGTCG 822
Db ATGAGTCTCCAGCTTGGCTCAGCTCTCCAGAGAGAGTGGTCCCTCTGTCGTCG 744
Qy TTGGAACCCAGGATC- AGGAGGATTTGGAGCCGTGAAGGAAGAAATGAGAGGATG 880
Db TTGGAACCCAGGATCAGGAGGATTTGGAGCCGTGAAGGAAGAAATGAGAGGATG 804
Qy GGGACCTT- GACCTGAACGGGAGTGTGTTG---TCGACAAACCGCTAGAAATGCCA 935
Db GGGACCTTGAACCTGAGCGGGGAGTGTGTTGTTGGTGGCACAACCCCGCTAGAAATGGCCC 864
Qy 936 AACCG 940
Db 865 AACCG 869

RESULT 13
BQ924736
LOCUS BQ924736
DEFINITION AGENCOURT_8853777 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6374110
5', mRNA sequence.
ACCESSION BQ924736
VERSION BQ924736.1 GI:22339767
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 938)
NIH-MGC http://imgc.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: sgabsr@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://imgc.llnl.gov
Plate: L1CM2552. row: i column: 23
High quality sequence stop: 629.
Location/Qualifiers

source 1. .938
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6374110"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

Query Match 26.7%; Score 714; DB 13; Length 938;
Best Local Similarity 89.7%; Pred. No. 1.3e-108;
Matches 874; Conservative 0; Mismatches 47; Indels 53; Gaps 8;

Qy 48 CAGCGGTTCCGGGTCGGCGGGGGGGGATGTAATCCCATGAGCGGCCGAGGA 107
Db 1 CAGCGGTTCCGGGTCGGCGGGGGGGGATGTAATCCCATGAGCGGCCGAGGA 60
Qy 108 AGCAACAGTCGCGCGCGCGCGAGCCCTGGGACGGCTCCTGCGTCTGGCGCGGAGGA 167
Db 61 AGCAACAGTCGCGCGCGCGCGAGCCCTGGGACGGCTCCTGCGTCTGGCGCGGAGGA 120
Qy 168 GGGGACCGCACTGCTCTCTGAGGAAGCGGAGTGGACCATCGGCGGAGAGAGGTTG 227
Db 121 GGGGACCGCACTGCTCTCTGAGGAAGCGGAGTGGACCATCGGCGGAGAGAGGTTG 180
Qy 228 CGACCTTCTCTCCCGAGCAATAAACTGGTCTCTGGAGATCACTGTAGAAATTTAGTGA 287
Db 181 CGACCTTCTCTCCCGAGCAATAAACTGGTCTCTGGAGATCACTGTAGAAATTTAGTGA 240
Qy 288 TGAATAATCAGGTGAGTGCATCTGGAAGATACAGCAACAGTGGAAACAGTGAACAA 347
Db 241 TGAATAATCAGGTGAGTGCATCTGGAAGATACAGCAACAGTGGAAACAGTGAACAA 300
Qy 348 GCTGAAGTGTGTTAAGAGCAGACATGCCCTTTACAGCTGGGATGTCATCTACTTGGT 407
Db 301 GCTGAAGTGTGTTAAGAGCAGACATGCCCTTTACAGCTGGGATGTCATCTACTTGGT 360
Qy 408 GTACAGGAAGATGAACCGGAACACAACTGGTGCATACCTCTATGAATCTTTAAGTGA 467
Db 361 GTACAGGAAGATGAACCGGAACACAACTGGTGCATACCTCTATGAATCTTTAAGTGA 420
Qy 468 GCAGGACATGACACAGAAATCTTTGAGCTTAACAGGAAATGTTCCATGGGACCAA 527
Db 421 GCAAGGACATGACACAGAAATCTTTT 445
Qy 528 AGATACCTCAGGTGACAGGCGAGGGCGGATGCCCGGGTCCCTCGTGGTCCGCC 587
Db 446 -GATACCTCAGGTGACAGGCGAGGGCGGATGCCCGGGTCCCTCGTGGTCCGCC 504
Qy 588 CGGCACCTCAGGTGCTTTGAGGAACACAGCCATCAACATCGACGTCAGACCTCTTCCC 647
Db 505 CGGCACCTCAGGTGCTTTGAGGAACACAGCCATCAACATCGACGTCAGACCTCTTCCC 564
Qy 648 CACAGCCTCGGCTCTTTCCACAGGAGCTTCTCTGAGGCGGAGAGCGTTCCTCCAGTTG 707
Db 565 CACAGCCTCGGCTCTTTCCACAGGAGCTTCTCTGAGGCGGAGAGCGTTCCTCCAGTTG 624
Qy 708 TGGGTCTGGGGTGGTGGCATCTCCCTTAAAGGAAGTGGTGGTGGTGGTGGTGGTGGT 766
Db 625 TGGGTCTGGGGTGGTGGCATCTCCCTTAAAGGAAGTGGTGGTGGTGGTGGTGGTGGT 684
Qy 767 AAGTCTCAGCTTTTGGCTCAGCTCTCCAGACAGAAAGACTGGCTCTCTTTCGTCGTTGG 826
Db 685 AAGTCTCAGCTTTTGGCTCAGCTCTCCAGACAGAAAGACTGGCTCTCTTTCGTCGTTGG 744

1870 GTACGGGAGACACCGTTCTGTGTTTCTGCTGTGCTGCGCAGCTTCCGTGAGCTGACC 1929
62 GTACGGGAGACACCGTTCTGTGTTTCTGCTGTGCTGCGCAGCTTCCGTGAGCTGACC 121
1930 TATCAGTATCGGAGAACATTCCTGCTCCGAGTTGCCAGTGGCGGTAAACATCCCGTCT 1989
122 TATCAGTATCGGAGAACATTCCTGCTCCGAGTTGCCAGTGGCGGTAAACATCCCGTCT 181
1990 GACTGCTACTGGGGCGGTAACTGCGCCACTCAGGTGAAAGCTCAACGCCATGAAATTC 2049
182 GACTGCTACTGGGGCGGTAACTGCGCCACTCAGGTGAAAGCTCAACGCCATGAAATTC 241
2050 AATCATATCTGTGACACAGCAAGGTTCAAAAATAAGCATCCAGAGGCCCTCAGAGAGCTT 2109
242 AATCATATCTGTGACACAGCAAGGTTCAAAAATAAGCATCCAGAGGCCCTCAGAGAGCTT 301
2110 TCAGCACTGGAGGTGAAGAGAGCGTGTGTTTAAATACAGACACAGCAAGCTCAAGGTGT 2169
302 TCAGCACTGGAGGTGAAGAGAGCGTGTGTTTAAATACAGACACAGCAAGCTCAAGGTGT 361
2170 TTTTCAGAGCCCTGAGGAGGAGCGAGGCTCTCCGACAGTGTCTCTGGGGTACTCT 2229
362 TTTTCAGAGCCCTGAGGAGGAGCGAGGCTCTCCGACAGTGTCTCTGGGGTACTCT 421
2230 TCTGTGAGAGCTTTTACCCCTCTGAGTGAGACCCCTCCCAAGAGCCCGGGGGCGGAGCC 2289
422 TCTGTGAGAGCTTTTACCCCTCTGAGTGAGACCCCTCCCAAGAGCCCGGGGGCGGAG-CC 478
2290 GCCTCTGCTGAGCGCTGGGAGAGGCTGCTGGTGGCATCAGCAGCAGAGAGAGAGCCCTT 2349
479 GCCTCTGCTGAGCGCTGGGAGAGGCTGCTGGTGGCATCAGCAGCAGAGAGAGAGCCCTT 538
2350 TCTGTAACTGCGCGCTGCCGCGAGAGGCGGAGGCTTTTGTCTCTTTGTACATTTTCCGA 2409
539 TCTGTAACTGCGCGCTGCCGCGAGAGGCGGAGGCTTTTGTCTCTTTGTACATTTTCCGA 598
2410 AACTAGCTAAAGCAGAGTCTGTTTTCAGG-AAAAGTTTCAGGAGAGAGAGGCAAGTT 2468
599 AACTAGCTAAAGCAGAGTCTGTTTTCAGGAGAGAGGCTTTTGTCTCTTTGTACATTTTCCGA 658
2469 TATCAAAAACATGTTTCAGGAGAGGAGCAGTAAAGTTTACAGCCTAC-AGGACGTACAC 2527
659 TATCAAAAACATGTTTCAGGAGAGGAGCAGTAAAGTTTACAGCCTTCAAGGAGGTACAC 717
2528 AATATCTGCTGCTGGGAGAGAGCAGTAAAGTTTATCTATTTTATTTATAGGTTGG 2587
718 AATATCTGCTGCTGGGAGAGAGCAGTAAAGTTTATCTATTTTATTTATCTAATA- ---GGTTGG 773
2588 TGCTTATCTTCTAATAGATTTTAAATGTCACAACTGTAGCACAATAAT 2637
774 GGCTTATCTTCTAATAGATTTTAAATGTCACAACTGTAGCACAATAAT 823

RESULT 15
BX409790 944 bp mRNA linear EST 13-MAY-2003
LOCUS BX409790 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF017119 5-PRIME, mRNA sequence.
ACCESSION BX409790
VERSION BX409790.1 GI:30656944
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 944)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

827 AA-CCCCAGGATCAGGAGATTTCGAGCCC-----GTCAAGAGAAATAGAGAGATG 880
745 AACCCCCAGATCAGGAGATTTCGAGCCC-----GTCAAGAGAAATAGAGAGATG 804
881 GGGACCTTTGACCTGAA-----CGGCGAGTTTGTGTCACAAACCGGTGAGAAATGCCCA 935
805 GGACCTTTGACCTGAAACCGGCGAGTTTGTGTCGACCAACCGGCTAGGAAATGGCCCA 864
936 AACCGTTCACAGGAG--CGTCAGAGCAGCGGCTGGGAG--CCAGACAAAGAT--CGAGGAG 990
865 AACCGTTCACAGGAGAACCCAGAGCAACCCGCTGGGAGCCCAAGATGGGAGAG 924
991 ACGCTGACATGCAT 1004
925 ACGCTGACATGCAT 938

RESULT 14
BG831289 824 bp mRNA linear EST 22-MAY-2001
LOCUS 602766120F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4908316 5',
DEFINITION mRNA sequence.
ACCESSION BG831289
VERSION BG831289.1 GI:14178876
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 824)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1814 row: 9 column: 05
High quality sequence stop: 745.
Location/Qualifiers
1. .824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4908316"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="pH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;
Site: 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library. |"

ORIGIN
Query Match 26.4%; Score 708.4; DB 12; Length 824;
Best Local Similarity 95.7%; Pred. No. 1.1e-107;
Matches 794; Conservative 0; Mismatches 26; Indels 10; Gaps 6;
1810 ATGTTGACCGAGAGCTCGGTCTCCAGCGGGAGTGTCTGCTGCTGATACAGA 1869
2 ATGTTGACCGAGAGCTCGGTCTCCAGCGGGAGTGTCTGCTGCTGATACAGA 61

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6792.x For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAF014ZC07_AFO1300_1&cluster=6792.x. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/InvitroGen> Corporation 1600 Paraday Avenue Genoscope sequence ID : CS0BAF014ZC07_AFO1300_1.

FEATURES

source
1..944
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF017Y119"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 26.4%; Score 706.2; DB 13; Length 944;
Best Local Similarity 89.7%; Pred. No. 2.5e-107;
Matches 849; Conservative 0; Mismatches 84; Indels 13; Gaps 8;

QY	1578	ACGGAGCGGAGCAGACCCGGCTGCGCCCTCAGCAGTGTGGGCTGCTGCTCAGCC	1637
Db	4	ACGGAGCGGAGCAGACCCGGCTGCGCCCT-AGCAGTGTGGGCTGCTGCTCAGCC	62
QY	1638	TTTCTCCCACTGTACTGGGGCTGCACCCGGACCGGCTGCTACGGCTGCTGGCCCGTT	1697
Db	63	TTTCTCCCACTGTACTGGGGCTGCACCCGGACCGGCTGCTACGGCTGCTGGCCCGTT	122
QY	1698	TTGTGAGCTCACTGGGTGACAAAGTGTCTGACGGCGGTGCTGACACACAGCTACGA	1757
Db	123	TTGTGAATCTGTAATGAAAGAAATATGTTATCTCAGTTCGGGCGACAGTGTCTTC	182
QY	1758	GTCCAGACATCTCA-----AGAAATTACCTGGCAACCCAGAGGTTTGACATGGA	1812
Db	183	CTCCGACATCATTTATAACAGAAATTAACCTGGCAACCCAGAGGTTTGACATGGA	242
QY	1813	TTGACCGAGAGCCTCGTGGCTCTCCAGCGGGAGTGTCTGCTGTCTGATTACAGAGTC	1872
Db	243	TTGACCGAGAGCCTCGTGGCTCTCCAGCGGGAGTGTCTGCTGTCTGATTACAGAGTC	302
QY	1873	ACGGGAGACACCTGTTGTGTACTGTGCTGCGCAGCTTCCGTCAGCTGACCTAT	1932
Db	303	ACGGGAGACACCTGTTGTGTACTGTGCTGCGCAGCTTCCGTCAGCTGACCTAT	362
QY	1933	CAGTATCGGCAGAACATTCCTGCTTCGGAGTTCCAGTGGCGGTAAACATCCCGTCTGAC	1992
Db	363	CAGTATCGGCAGAACATTCCTGCTTCGGAGTTCCAGTGGCGGTAAACATCCCGTCTGAC	422
QY	1993	TGCTACTGGGGCGGTAACTGCCGACTCAGGTGAAAGCTCACCAGCCCATGAAATTCAT	2052
Db	423	TGCTACTGGGGCGGTAACTGCCGACTCAGGTGAAAGCTCACCAGCCCATGAAATTCAT	482
QY	2053	CATATCTGTGAACAGACAGGTTCAAAACTAAAGCATCCAGAGGCCCTGAGCAGCTTCA	2112
Db	483	CATATCTGTGAACAGACAGGTTCAAAACTAAAGCATCCAGAGGCCCTGAGCAGCTTCA	542
QY	2113	GCATGAGGTGAAGAGAGCGTGTGTTTAAATACAGAGACAGCAGCTCAAGTGTGTTT	2172
Db	543	GCATGAGGTGAAGAGAGCGTGTGTTTAAATACAGAGACAGCAGCTCAAGTGTGTTT	602
QY	2173	CACAGCCCCCTGAGGGAGGAGCGAGGGTCTCCGACAGGTGCTCTGGGGTGACTCTTCT	2232
Db	603	CACAGCCCCCTGAGGGAGGAGCGAGGGTCTCCGACAGGTGCTCTGGGGTGACTCTTCT	662
QY	2233	GTGGAGCTTTTACCCTCTGAGTGAGACCTCCCCAGAGCCCCGGGGCGGAGCCCGCC	2292

Search completed: May 14, 2004, 20:06:21
Job time : 10045 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 20:10:20 ; Search time 16426 Seconds
(without alignments)
7069.036 Million cell updates/sec

Title: US-10-048-046-1
Perfect score: 2679
Sequence: 1 aggaattcgccagggccg.....acaaaaaaaaaaaaaaaaa 2679

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.btg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2679	100.0	2679	9	AF170724	AF170724 Homo sapi
2	2133	79.6	2639	6	AX877194	AX877194 Sequence
3	2133	79.6	2639	6	BD156534	BD156534 Primer fo
4	2133	79.6	2639	9	AK027687	AK027687 Homo sapi
5	2031	75.8	3189	9	BC012072	BC012072 Homo sapi
6	1818	60.4	3181	6	AX405642	AX405642 Sequence
7	1497	55.9	2297	9	HSW802304	AL137561 Homo sapi
8	1465	54.7	3138	6	AX877222	AX877222 Sequence
9	1465	54.7	3138	6	BD156548	BD156548 Primer fo
10	1465	54.7	3138	9	AK001658	AK001658 Homo sapi
11	944	35.2	2448	6	AX834949	AX834949 Sequence
12	944	35.2	2448	9	AK097671	AK097671 Homo sapi
13	930	34.7	6235	6	BD183291	BD183291 Novel gen
14	909	33.9	2257	9	AK090948	AK090948 Homo sapi
15	633	23.6	2186	6	AX713476	AX713476 Sequence
16	633	23.6	2186	9	AX054917	AK054917 Homo sapi
17	478	17.8	816	6	AX868758	AX868758 Sequence
18	478	17.8	816	6	BD148820	BD148820 Primer fo
19	468	17.5	824	6	AX868776	AX868776 Sequence
20	468	17.5	824	6	BD148838	BD148838 Primer fo
21	426	15.9	518	6	AX873789	AX873789 Sequence
22	426	15.9	518	6	BD153851	BD153851 Primer fo
23	386	14.4	181012	9	AC127070	AC127070 Homo sapi
24	386	14.4	181438	2	AC023047	AC023047 Homo sapi
25	218	8.1	181438	2	AC023047	AC023047 Homo sapi
26	78	2.9	617	6	AX385952	AX385952 Sequence
27	38	1.4	158897	2	AC118260	AC118260 Mus muscu
28	37	1.4	3194	10	BC049792	BC049792 Mus muscu
29	37	1.4	321708	2	AC087142	AC087142 Mus muscu
30	37	1.4	338116	2	AC087159	AC087159 Mus muscu
31	34	1.3	202792	2	AC125898	AC125898 Rattus no
32	32	1.2	601	11	BV002855	AC124951 Medicago
33	29	1.1	125396	2	AC124951	AC124951 Medicago
34	28	1.0	146516	2	AC079981	AC079981 Homo sapi
35	28	1.0	218226	2	AC126634	AC126634 Rattus no
36	28	1.0	228987	2	AC111838	AC111838 Rattus no
37	28	1.0	274796	2	AC097665	AC097665 Rattus no
38	27	1.0	167166	2	AC117568	AC117568 Mus muscu
39	27	1.0	167904	2	AC115904	AC115904 Mus muscu
40	27	1.0	217842	2	AL929073	AL929073 Mus muscu
41	27	1.0	230221	10	AC123851	AC123851 Mus muscu
42	27	1.0	234393	10	AL772311	AL772311 Mouse DNA
43	27	1.0	236825	2	AC140355	AC140355 Mus muscu
44	27	1.0	239783	2	AC102564	AC102564 Mus muscu
45	27	1.0	266632	2	AC129762	AC129762 Rattus no

ALIGNMENTS

RESULT 1
AF170724
LOCUS AF170724 Homo sapiens cell cycle checkpoint protein CHFR mRNA, complete cds.
DEFINITION AF170724 Homo sapiens cell cycle checkpoint protein CHFR mRNA, complete cds.
ACCESSION AF170724
VERSION AF170724.1 GI:9651169
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2679)
Scolnick,D.M. and Halazonetis,T.D.
TITLE Chfr defines a mitotic stress checkpoint that delays entry into metaphase

Pred. No. is the number of results predicted by chance to have a

```

JOURNAL Nature 406 (5794), 430-435 (2000)
MEDLINE 20388685
PUBMED 10935642
REFERENCE 2 (bases 1 to 2679)
AUTHORS Halazonetis,T.D. and Scolnick,D.M.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1999) Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA
FEATURES
    source
        Location/Qualifiers
            1..2679
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
            91..2085
                /note="contains FHA domain and ring finger"
            /codon_start=1
            /product="cell cycle checkpoint protein CHFR"
            /protein_id="AAF91084.1"
            /db_xref="GI:9651170"
            /translation="MEPEBGGKSPPPQWGRLLRLGAEGEPHVLLRKREWTIGRR
            CGLDSPSNKLVSCHDRIIVDEKSGVQLEDTSITGVINKLKVKKQTCPLQGDV
            ILYLRKNEPHNCHVIVYESLSKQWQTESPEAKENVFVKTDKTSAGAGRADPR
            VPPSPQVCFCEPQPSSTSDLPFTASSTSEPPAGRRSSCSGGGGISPKGS
            GVSASDEVSFASALPDKRTASFSLEPQQDELEPKVKRWGGDGLDLGQLLVAQ
            PRNAQTVHEVRAAGKPKMEETLTCIIICOLLHDVCSPQCMCHTFCACYSGMWE
            RSSLCPTCRCPVERIKCNHILNVLVAYLIQHPDKSRSEEDVQSMARKNTIOMDLQ
            KYRSEDESGESDEILELSDVDSSESDISQPVVCRQCPSEYRQAQPPCHCAPGE
            PQACALGDAPFSTSVLITVAQVYVCPQCSHALCTCCPQPMPEARAREQDPFVAQ
            QCVLCVQFCHLYGCTRTCGCLAPFCLELNGDKLGDVLNNNSYESDILKNLYAT
            RGLTKMNLTESIVALQGRVFLSDYRVTDGLTVLCYCGLRSFRLTYVQRNPAPSE
            LPVAVTSRPDCYMGRCNRTQVKAHAMKFNHICEQTRFKN"
    CDS
        100.0%; Score 2679; DB 9; Length 2679;
        Query Match
        Best Local Similarity 100.0%; Pred. No. 0;
        Matches 2679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

ORIGIN

Query Match	100.0%; Score 2679; DB 9; Length 2679;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2679; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AAGAAATTCGGACAGCGCGCAATGTCCTTTGACAGCGCGCGCGAGCCGGTTCGGG 60
DB	1 AAGAAATTCGGACAGCGCGCAATGTCCTTTGACAGCGCGCGCGAGCCGGTTCGGG 60
QY	61 GTTCGGCGCGGGCGGGGATGTGAATCCCGATGGAGCGGCGCGAGAGCGCAAGCAGTGC 120
DB	61 GTTCGGCGCGGGCGGGGATGTGAATCCCGATGGAGCGGCGCGAGAGCGCAAGCAGTGC 120
QY	121 CGCGCGCGCAGCCCTGGGGACGGCTCCTCGGTCCTGGGCGCGGAGGCGGAGCCGCAC 180
DB	121 CGCGCGCGCAGCCCTGGGGACGGCTCCTCGGTCCTGGGCGCGGAGGCGGAGCCGCAC 180
QY	181 GTCTCCTCTGAGGAAGCGGGAGTGGACCATCGGCGCGAGACGAGGTTGCGACCTTTCCCTTC 240
DB	181 GTCTCCTCTGAGGAAGCGGGAGTGGACCATCGGCGCGGAGACGAGGTTGCGACCTTTCCCTTC 240
QY	241 CCGAGCAATAAATCTGGTCTCTGGAGATCATCTGTAGAAATTGTAGTGGATGAAAAATCAGGT 300
DB	241 CCGAGCAATAAATCTGGTCTCTGGAGATCATCTGTAGAAATTGTAGTGGATGAAAAATCAGGT 300
QY	301 CAGGTGACACTGGAGAGATACCAGACCAAGTGGACAGTGAATTAACAAGCTGAAGGTGTT 360
DB	301 CAGGTGACACTGGAGAGATACCAGACCAAGTGGACAGTGAATTAACAAGCTGAAGGTGTT 360
QY	361 AAGAAAGCAGACATGCCCTTTTACAGACTGGGGATGTCACTTACTTGGTGTACAGGAAAGAT 420
DB	361 AAGAAAGCAGACATGCCCTTTTACAGACTGGGGATGTCACTTACTTGGTGTACAGGAAAGAT 420
QY	421 GAACCGGAACACAACGTTGGCATACCTCTATGATCTTTAAGTGAAGAAGCAAGGCATGCA 480
DB	421 GAACCGGAACACAACGTTGGCATACCTCTATGATCTTTAAGTGAAGAAGCAAGGCATGCA 480
QY	481 CAAGAATCCTTTGAAGCTTAAACAAGGAAAAATGTGTTTCCATGGGACCAAGATACCTCAGGT 540
DB	481 CAAGAATCCTTTGAAGCTTAAACAAGGAAAAATGTGTTTCCATGGGACCAAGATACCTCAGGT 540

541	GCAGGTGCAGGGCGAGGGGCGCATCCCGGGGTCCCTCCGTCTGTCCGCGCCACTCAGGTG	600	QY
541	GCAGGTGCAGGGCGAGGGGCGCATCCCGGGGTCCCTCCGTCTGTCCGCGCCACTCAGGTG	600	DB
601	TGCTTTGAGGAACACAGGCCATCAACATCGACGTGACAGCTCTTCCCCACAGCTCGGCC	660	QY
601	TGCTTTGAGGAACACAGGCCATCAACATCGACGTGACAGCTCTTCCCCACAGCTCGGCC	660	DB
661	TCCTTCCACGGAGCCTTCTCCTCGAGGGCGAGAGCGTTCCCTCAGATTGTGGGCTCTGGGGT	720	QY
661	TCCTTCCACGGAGCCTTCTCCTCGAGGGCGAGAGCGTTCCCTCAGATTGTGGGCTCTGGGGT	720	DB
721	GGTGGCATCTCCCTTAAGGAAGTGGTCCCTCTGTGGCAAGTGATGAAGTCTCCAGCTTT	780	QY
721	GGTGGCATCTCCCTTAAGGAAGTGGTCCCTCTGTGGCAAGTGATGAAGTCTCCAGCTTT	780	DB
781	GGCTCAGCTCTCCACAGACAGAAAGACTGCGTCTCTTTTCGTCTGTGGAAACCCACAGATCAG	840	QY
781	GGCTCAGCTCTCCACAGACAGAAAGACTGCGTCTCTTTTCGTCTGTGGAAACCCACAGATCAG	840	DB
841	GAGGATTTGGAGCCCGTGAAGAAATAATGAGAGGATGCGGACCTTCGACTGACAGCGG	900	QY
841	GAGGATTTGGAGCCCGTGAAGAAATAATGAGAGGATGCGGACCTTCGACTGACAGCGG	900	DB
901	CAGTTGTTGGTGCACAAACCGGTGAGAAATGCCAAAACCGTCCACGAGGACGTGACAGCA	960	QY
901	CAGTTGTTGGTGCACAAACCGGTGAGAAATGCCAAAACCGTCCACGAGGACGTGACAGCA	960	DB
961	GGCGTGGGAAGCCAGACAAGATGAGGAGAGCGTGACATGATCATCTGCCAGGACCTG	1020	QY
961	GGCGTGGGAAGCCAGACAAGATGAGGAGAGCGTGACATGATCATCTGCCAGGACCTG	1020	DB
1021	CTGCACGACTGCGTGAGTTTGAGCGCCCTGCATGCAACGTTCTGCGCGGCTTCTACTCG	1080	QY
1021	CTGCACGACTGCGTGAGTTTGAGCGCCCTGCATGCAACGTTCTGCGCGGCTTCTACTCG	1080	DB
1081	GGCTGATGGAGCGTCTGCTCTGTGCTCTACCTGCGCTGTCTCCCTGGAGCGGATCTGT	1140	QY
1081	GGCTGATGGAGCGTCTGCTCTGTGCTCTACCTGCGCTGTCTCCCTGGAGCGGATCTGT	1140	DB
1141	AAAAACCAATCCTCAACAACTCTGTGAAGCATACCTCATCCAGCATCCAGACAAGAT	1200	QY
1141	AAAAACCAATCCTCAACAACTCTGTGAAGCATACCTCATCCAGCATCCAGACAAGAT	1200	DB
1201	CGCAGTGAAGAGATGTGCAAAGTATGGATGCCAGGAATAAATCACTCAAGACATGCTG	1260	QY
1201	CGCAGTGAAGAGATGTGCAAAGTATGGATGCCAGGAATAAATCACTCAAGACATGCTG	1260	DB
1261	CAGCCCAAAGTCAGGCGGTCTTTTCTGTGAAGAGGAGTTTCAGAGGACCTGCTGGAG	1320	QY
1261	CAGCCCAAAGTCAGGCGGTCTTTTCTGTGAAGAGGAGTTTCAGAGGACCTGCTGGAG	1320	DB
1321	CTGTGAGCGTTGACAGTCAGTCTCCTCAGACATTAGCCAGCCATACGTCTGTCGCCGCGAG	1380	QY
1321	CTGTGAGCGTTGACAGTCAGTCTCCTCAGACATTAGCCAGCCATACGTCTGTCGCCGCGAG	1380	DB
1381	TGTCCTGAGTACAGAAAGCAGCGGCGCAGCTCCCCATGCCCCAGCACCCGAGGGCGAG	1440	QY
1381	TGTCCTGAGTACAGAAAGCAGCGGCGCAGCTCCCCATGCCCCAGCACCCGAGGGCGAG	1440	DB
1441	CCAGAGCCCCACAGGCCCTGGGGATGACACCTCCACGTCGTTCAGCTGACGACAGCA	1500	QY
1441	CCAGAGCCCCACAGGCCCTGGGGATGACACCTCCACGTCGTTCAGCTGACGACAGCA	1500	DB
1501	GTCACGATTACGTGTGCCCTCTGCAAGGAAGCCACGCCCTGTGCACTGCTGCTCCAG	1560	QY
1501	GTCACGATTACGTGTGCCCTCTGCAAGGAAGCCACGCCCTGTGCACTGCTGCTCCAG	1560	DB
1561	CCCATGCCGACCGAGAGCGGAGCGCGAGCAGGACCCGGCTGTCTGCCCTCAGCAGTGT	1620	QY
1561	CCCATGCCGACCGAGAGCGGAGCGCGAGCAGGACCCGGCTGTCTGCCCTCAGCAGTGT	1620	DB
1621	GGCGTCTGCTCAGCCTTTCTGCAACCTGTACTGGGGGTGCACCCCGGACGGGCTGCTAC	1680	QY

1621 GCAGTCTGCTGAGCGCTTTCTGCCACCTGTACTGGGGCTGCACCGGACCGGCTGCTAC 1680
1681 GCGTCTGCTGCGCCGCTTTGTGAGCTCAACCTGGGTGACAGTGTCTGACGCGCTGCTG 1740
1681 GCGTCTGCTGCGCCGCTTTGTGAGCTCAACCTGGGTGACAGTGTCTGACGCGCTGCTG 1740
1741 AACACAAACAGCTTACGAGTACAGATCCCTGAAGAATTACCTGGCAACACAGAGTTTGACA 1800
1741 AACACAAACAGCTTACGAGTACAGATCCCTGAAGAATTACCTGGCAACACAGAGTTTGACA 1800
1801 TGGAAAAACATGTTGACACAGAGCTGCTGGCTCTCCAGCGGGAGTGTCTTCTGCTGCT 1860
1801 TGGAAAAACATGTTGACACAGAGCTGCTGGCTCTCCAGCGGGAGTGTCTTCTGCTGCT 1860
1861 GATTACAGAGTACAGGAGACACCGTCTGCTGTTACTGCTGCTGGCTGCGAGCTTCCGT 1920
1861 GATTACAGAGTACAGGAGACACCGTCTGCTGTTACTGCTGCTGGCTGCGAGCTTCCGT 1920
1921 GAGCTACCTATCAGTATCGGCAAGACATCTCTGCTTCCGAGTGTGCAAGTGGCGGTAAACA 1980
1921 GAGCTACCTATCAGTATCGGCAAGACATCTCTGCTTCCGAGTGTGCAAGTGGCGGTAAACA 1980
1981 TCCCGTCTGACTGCTACTGGGCGCTTAACTGCGGCTCACTGAGTGAAGCTCACCACGCC 2040
1981 TCCCGTCTGACTGCTACTGGGCGCTTAACTGCGGCTCACTGAGTGAAGCTCACCACGCC 2040
2041 ATGAATTTCAATCATATCTGTGAACAGACAGGTTTCAAAAACTAAGCATCCAGAGCCCT 2100
2041 ATGAATTTCAATCATATCTGTGAACAGACAGGTTTCAAAAACTAAGCATCCAGAGCCCT 2100
2101 GAGCAGCTTCAGCTGAGGCTGAAGAGGCTGTTTAAAAATACAGACAGAGCAG 2160
2101 GAGCAGCTTCAGCTGAGGCTGAAGAGGCTGTTTAAAAATACAGACAGAGCAGCAG 2160
2161 TCAAGTGTTTTCACAGCCCCCTGAGGAGGAGCAGAGGCTCTCCGACAGTGTCTTGG 2220
2161 TCAAGTGTTTTCACAGCCCCCTGAGGAGGAGCAGAGGCTCTCCGACAGTGTCTTGG 2220
2221 GGTGACTCTTCTGAGCTTTTACCTCTGAGTGAGACCTTCCCGAGGCCCCGGGG 2280
2221 GGTGACTCTTCTGAGCTTTTACCTCTGAGTGAGACCTTCCCGAGGCCCCGGGG 2280
2281 CCGCAGCCGCCCTCTGCTGAGCTGCGGAGGCTCTGCTGAGTGGATCAGCAGCAGAGA 2340
2281 CCGCAGCCGCCCTCTGCTGAGCTGCGGAGGCTCTGCTGAGTGGATCAGCAGCAGAGA 2340
2341 CGAAGCTTTCTGTAACTGCGGCGCTTCCCGAGGAGGAGGCTTCTCTTTTGTAC 2400
2341 CGAAGCTTTCTGTAACTGCGGCGCTTCCCGAGGAGGAGGCTTCTCTTTTGTAC 2400
2401 ATTTTCGAACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAGTTCAGAGGAGAG 2460
2401 ATTTTCGAACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAGTTCAGAGGAGAG 2460
2461 GGCAAGTTTATCAAAACATTTTTCAGGAGAAAGGAGCATTAAGTTTACAGCTACAGGA 2520
2461 GGCAAGTTTATCAAAACATTTTTCAGGAGAAAGGAGCATTAAGTTTACAGCTACAGGA 2520
2521 CGTACACATATCTGCTGCTGGGAAACCAAGAGATTTTCTATTTTATTTTATA 2580
2521 CGTACACATATCTGCTGCTGGGAAACCAAGAGATTTTCTATTTTATTTTATA 2580
2581 GGTTCGTGCTTATCTTCTAATAAGATTTTAAATGTTCACAACTGTAGCACAATAATATA 2640
2581 GGTTCGTGCTTATCTTCTAATAAGATTTTAAATGTTCACAACTGTAGCACAATAATATA 2640
2641 ATTTTAAATTTACAAATTTGACAAAAAATTTTAAAAA 2679
2641 ATTTTAAATTTACAAATTTGACAAAAAATTTTAAAAA 2679

RESULT 2

AX877194
LOCUS AX877194 2639 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 12099 from Patent EP1074617.
ACCESSION AX877194
VERSION AX877194.1 GI:40031930
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 12099 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1..2639
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
79..2037
/note="unnamed protein product"
/codon_start=1
/protein_id="CA889614.1"
/db_xref="GI:40031931"
/translation="MERPEEGKQSPPPQWGRLLRLGAERGEHVLRLKREWTIGRR
GCDLSPSNKLVSGDHCRIVVDEKSGQVLTEDSTGIVINKLVKVKQTCPIQTGV
IYLVRNKEPEHNVADLYESLKSQGTQESFDTSGAGARGADRPVPSPATQVCF
EEOPTSTSDLPFTASSTEPSRAGRSRSGSGGGISPKSGSPVASDEVSVF
ASALPDKTASFSLEPQDOEDLEPVKKRNGDGLDLNQLVAQPRNAQTVEHD
RAAGKPKMEETLTGIIICODLLHDCVSLQPCVHTFCAACYSYGMWSSSLCPTCRPV
ERIKKHILNVLVEAYLIQHPDKSRSDPQSDMDANKITQDMLQPKVRSFSDBEGS
SEDLLESLDVSSESSDISOPVVCROCPVEYRORAPHPGAPGEPGAPALGDAPS
TSVSLTTAVQDVCPLQSHALCTCCQPPMDRAEREQDPVPAQCAVCLQPFCHL
YMGCTRTGCGCLAPFCLELNDKCLDGLNNNSYESDILKNLYLATLGLTWKNLWTS
LVALQRGVFLLSLYRTGTVLCVCGGLRSFRELTYQYQNI PASLPLFAVTSRPDCY
WGRNCRTOVKAHAMKENHICEQTRFKN"

Query Match 79.6%; Score 2133; DB 6; Length 2639;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 529 GATACCTCAGGTGAGGTGAGGGCGAGGGCGCGATCCCGGGTCCCTCGGTCTCGGCC 588
DB 481 GATACCTCAGGTGAGGTGAGGGCGAGGGCGCGATCCCGGGTCCCTCGGTCTCGGCC 540
QY 589 GCGACTCAGGTGCTTTGAGGAACCAAGCCATCAATCGAGCTCAGACCTCTTCCC 648
DB 541 GCGACTCAGGTGCTTTGAGGAACCAAGCCATCAATCGAGCTCAGACCTCTTCCC 600
QY 649 ACAGCTCGGCTCTTCCAGGAGCTTCTCCTCGAGGGGAGAGCGTTCCTCAGTTGT 708
DB 601 ACAGCTCGGCTCTTCCAGGAGCTTCTCCTCGAGGGGAGAGCGTTCCTCAGTTGT 660
QY 709 GGGTCTGGGGTGTGTGGCATCTCCCTAAAGAAAGTGTCTCTGTGGCAGATGAA 768
DB 661 GGGTCTGGGGTGTGTGGCATCTCCCTTAAAGAAAGTGTCTCTGTGGCAGATGAA 720
QY 769 GTCTCCAGCTTGTCTCAGCTCTCCAGAGAGTGTCTCTCTCTTCTCTCTTGGAA 828
DB 721 GTCTCCAGCTTGTCTCAGCTCTCCAGAGAGTGTCTCTCTCTTCTCTCTTGGAA 780
QY 829 CCCAGGATCAGGAGGATTTGGAGCCCGTGAAGAAATGAGAGGAGATGGGACCTT 888
DB 781 CCCAGGATCAGGAGGATTTGGAGCCCGTGAAGAAATGAGAGGAGATGGGACCTT 840
QY 889 GACCTGAACGGGAGTGTCTGTGTCGACCAACCGGTGAGAAATGCCAAACCGTCCACGAG 948
DB 841 GACCTGAACGGGAGTGTCTGTGTCGACCAACCGGTGAGAAATGCCAAACCGTCCACGAG 900
QY 949 GACCTCAGAGCAGCGCTGGGAAGCCAGACAAAGATGGAGAGCGCTGACATGATCATC 1008

FEATURES	PT	CDS	Location/Qualifiers		Query Match	Best Local Similarity	Matches 2133; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
			Location/Qualifiers	(79)..(2034).							
Source			1..2639								
			/organism="Homo sapiens"								
			/mol_type="genomic DNA"								
			/db_xref="taxon:9606"								
ORIGIN											
Query Match 79.6%; Score 2133; DB 6; Length 2639;											
Best Local Similarity 100.0%; Pred. No. 0;											
Matches 2133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											
QY	529	GATACCTCAGGTGCGAGTGCAGGCGGAGGCGGATCCCGGGTCCCTCCGTCGTCGCC	588								
DB	481	GATACCTCAGGTGCGAGTGCAGGCGGAGGCGGATCCCGGGTCCCTCCGTCGTCGCC	540								
QY	589	GCCACTCAGGTGCTTTGAGGAACACAGCCATCAACATGACGTGACAGCTCTTCC	648								
DB	541	GCCACTCAGGTGCTTTGAGGAACACAGCCATCAACATGACGTGACAGCTCTTCC	600								
QY	649	ACAGCTCGGCTCTTCCAGGAGCTTCTCTGAGGCGGAGGCTTCTCCAGTGT	708								
DB	601	ACAGCTCGGCTCTTCCAGGAGCTTCTCTGAGGCGGAGGCTTCTCCAGTGT	660								
QY	709	GGTCTGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAGTATGAA	768								
DB	661	GGTCTGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAGTATGAA	720								
QY	769	GTCTCCAGCTTTGCTCCTCAGCTCTCCAGACAGAAAGACGTGCTCTTTGCTGGTGA	828								
DB	721	GTCTCCAGCTTTGCTCCTCAGCTCTCCAGACAGAAAGACGTGCTCTTTGCTGGTGA	780								
QY	829	CCCCAGGATCAGGAGATTTGAGCCCGTGAAGAGAAATCAGAGGAGATGGGACCTT	888								
DB	781	CCCCAGGATCAGGAGATTTGAGCCCGTGAAGAGAAATCAGAGGAGATGGGACCTT	840								
QY	889	GACCTGAACGGGCGAGTTGTGTGCAACACCGGTGAGAAATGCCAAACCGTCCACGAG	948								
DB	841	GACCTGAACGGGCGAGTTGTGTGCAACACCGGTGAGAAATGCCAAACCGTCCACGAG	900								
QY	949	GAGCTCAGAGCAGCGCTGGGAGGAGCAGACAGAGTGGAGGACCTGCATGATCATC	1008								
DB	901	GAGCTCAGAGCAGCGCTGGGAGGAGCAGACAGAGTGGAGGACCTGCATGATCATC	960								
QY	1009	TGCCAGGACCTGTGCACGACTGCGTAGTGTTCAGACCTGATGCACACGTTCTCGCG	1068								
DB	961	TGCCAGGACCTGTGCACGACTGCGTAGTGTTCAGACCTGATGCACACGTTCTCGCG	1020								
QY	1069	GCTTGCTACTCGGGCTGGATGGAGCGCTCGTCTGTGTCTTACCTGCGCTGTCCGCTG	1128								
DB	1021	GCTTGCTACTCGGGCTGGATGGAGCGCTCGTCTGTGTCTTACCTGCGCTGTCCGCTG	1080								
QY	1129	GAGCGGATCTGTAAAAACCAATCTCTCAACACCTCGTGGAGGATACCTCATCCAGCAT	1188								
DB	1081	GAGCGGATCTGTAAAAACCAATCTCTCAACACCTCGTGGAGGATACCTCATCCAGCAT	1140								
QY	1189	CCAGAACAAGTGCAGTGAAGAAGTGTGCAAGATATGGATGGCCAGGATAAATCACT	1248								
DB	1141	CCAGAACAAGTGCAGTGAAGAAGTGTGCAAGATATGGATGGCCAGGATAAATCACT	1200								
QY	1249	CAAGACATCTGCAGCCCAAGTCAAGCTCTTTTCTGATGAAGAAGGAGTTTCAGAG	1308								
DB	1201	CAAGACATCTGCAGCCCAAGTCAAGCTCTTTTCTGATGAAGAAGGAGTTTCAGAG	1260								
QY	1309	GACCTGCTGGAGTGTGACAGCTGTGACGTGAGTCTCTCAGACATTTAGCCAGCCATCGTC	1368								
DB	1261	GACCTGCTGGAGTGTGACAGCTGTGACGTGAGTCTCTCAGACATTTAGCCAGCCATCGTC	1320								
QY	1369	GTGTGCGGCGAGTCTCTGAGTACAGAGCGGCGGCGAGCTCCCTCCAGTCCAGCA	1428								
DB	1321	GTGTGCGGCGAGTCTCTGAGTACAGAGCGGCGGCGAGCTCCCTCCAGTCCAGCA	1380								

QY	1429	CCGAGGGCGAGCCAGGAGCCCCACAGGCCCTGGGGATGCACCTCCACGTCGTCAGC	1488								
DB	1381	CCGAGGGCGAGCCAGGAGCCCCACAGGCCCTGGGGATGCACCTCCACGTCGTCAGC	1440								
QY	1489	CTGACGACAGCAGTCCAGGATTTACGTGTGCTCTCTGCAAGAAAGCCACGCTGTGACC	1548								
DB	1441	CTGACGACAGCAGTCCAGGATTTACGTGTGCTCTCTGCAAGAAAGCCACGCTGTGACC	1500								
QY	1549	TGCTGCTTCCAGGCCATGCGGACCGGAGCGGAGCGGCGAGCAGGACCCGCGTGTGCGC	1608								
DB	1501	TGCTGCTTCCAGGCCATGCGGACCGGAGCGGAGCGGCGAGCAGGACCCGCGTGTGCGC	1560								
QY	1609	CCTCAGCAGTGTGCGGTCTGCTGACGCTTTTCTGCGACCTGTACTGGGCTGCACCCGG	1668								
DB	1561	CCTCAGCAGTGTGCGGTCTGCTGACGCTTTTCTGCGACCTGTACTGGGCTGCACCCGG	1620								
QY	1669	ACGGCTCTCAGCGTCTGCGGCCCGTCTTGTGAGCTCAACCTGGGTGACAGTGTCTG	1728								
DB	1621	ACGGCTCTCAGCGTCTGCGGCCCGTCTTGTGAGCTCAACCTGGGTGACAGTGTCTG	1680								
QY	1729	GACGGCTGTGTAACAACAACAGCTACAGTCAAGACATCTCTGAAGAATTACCTGGCAACC	1788								
DB	1681	GACGGCTGTGTAACAACAACAGCTACAGTCAAGACATCTCTGAAGAATTACCTGGCAACC	1740								
QY	1789	AGAGTTTGAATGGAAGAAAACATGTTGACGAGAGCTCTGCTGCTCTCCAGCGGGAGTG	1848								
DB	1741	AGAGTTTGAATGGAAGAAAACATGTTGACGAGAGCTCTGCTGCTCTCCAGCGGGAGTG	1800								
QY	1849	TTTCTGCTGTCTGATTAACAGATCAAGGAGACACGCTTCTGTGTTACTGCTGTGGCTG	1908								
DB	1801	TTTCTGCTGTCTGATTAACAGATCAAGGAGACACGCTTCTGTGTTACTGCTGTGGCTG	1860								
QY	1909	CGCAGCTTCCGTGAGCTGACCTATCATCTGCGAGAACATCTCTGCTCCGAGTTGCCA	1968								
DB	1861	CGCAGCTTCCGTGAGCTGACCTATCATCTGCGAGAACATCTCTGCTCCGAGTTGCCA	1920								
QY	1969	GTGGCCGTAAATCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2028								
DB	1921	GTGGCCGTAAATCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980								
QY	2029	GCTCACCACGCGCATGAAATTCATATCTGTGAAACAGACAAAGTTCAAAAATAAGCA	2088								
DB	1981	GCTCACCACGCGCATGAAATTCATATCTGTGAAACAGACAAAGTTCAAAAATAAGCA	2040								
QY	2089	TCCAGAGCGCTGAGCAGCTTTTCAAGCTGAGAGTGAAGAGGCTGTTTAAATATA	2148								
DB	2041	TCCAGAGCGCTGAGCAGCTTTTCAAGCTGAGAGTGAAGAGGCTGTTTAAATATA	2100								
QY	2149	GAGACAGCAGCTCAAGGCTGTTTTCAGAGCCCTGAGGGAAGGAGCGCAGGCTCTCCGA	2208								
DB	2101	GAGACAGCAGCTCAAGGCTGTTTTCAGAGCCCTGAGGGAAGGAGCGCAGGCTCTCCGA	2160								
QY	2209	CAGTGTCTCTGGGTGACTCTTCTGTGGAGCTTTTAACTCTGAGTGAAGACCTCCCA	2268								
DB	2161	CAGTGTCTCTGGGTGACTCTTCTGTGGAGCTTTTAACTCTGAGTGAAGACCTCCCA	2220								
QY	2269	GAGCCCGGGGGCGCAGCGCCCTCTGCTGAGCGCTGGGCGAGGCTCGTGTGGCAT	2328								
DB	2221	GAGCCCGGGGGCGCAGCGCCCTCTGCTGAGCGCTGGGCGAGGCTCGTGTGGCAT	2280								
QY	2329	CAGCAGCAGAGAGCAAGCCTTTCTGTAAATGCGCGCTCCCGCCGAGAGGGGAGTTT	2388								
DB	2281	CAGCAGCAGAGAGCAAGCCTTTCTGTAAATGCGCGCTCCCGCCGAGAGGGGAGTTT	2340								
QY	2389	GCTCTTTTGTACATTTTCCGAACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAAGTT	2448								
DB	2341	GCTCTTTTGTACATTTTCCGAACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAAGTT	2400								
QY	2449	TCAAAGGAGAGGCGCAAGTTTATCAAAAACATTTTTCAGGAGAGGAGGAGTAAAGTTTA	2508								
DB	2401	TCAAAGGAGAGGCGCAAGTTTATCAAAAACATTTTTCAGGAGAGGAGGAGTAAAGTTTA	2460								
QY	2509	CAGCTACAGAGCTACACATATCTCTGCTCTGGGAAAACCCACAGCATTTTATCTATT	2568								

Db 2461 CAGCCTACAGGACGTACACATATCTCTGCTGCTGGGAACACACAGCATTTATCTATTT 2520
Qy 2569 TTTATTTTAAATAGGTTTGGCTTATCTCTTAATAGATTAATATGTCACAACTGTAGC 2628
Db 2521 TTTATTTTAAATAGGTTTGGCTTATCTCTTAATAGATTAATATGTCACAACTGTAGC 2580
Qy 2629 ACAATATATATATTTATATATTTACAAATTTGAC 2661
Db 2581 ACAATATATATATTTATATATTTACAAATTTGAC 2613

RESULT 4

AK027687

LOCUS

DEFINITION

AK027687

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

Location/Qualifiers

1. .2639

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RP4000455"

/cell_line="NT2"

/cell_type="teratocarcinoma"

/clone_lib="NT2RP4"

/notes="Cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction."

79. .2037

/notes="unnamed protein product"

/codon_start=1

/protein_id="BAB55297.1"

/db_xref="GI:14042553"

/translation="MERPEGKQSPPPQWGLLLGLABEGEPHVLKRWTGRRR

GCGLSPFNKLVSQDHRIVDEKSGQVLTEDTSGTVINKLVKKQTCPLQTDV

ILVYKRNPEHNVAYLSEKQMTQESFTSGAGRGADPRVPPSPATQVCF

EPQPTSTSLPFTASSTFSPAGRSRSSCGSGGSGISPKSGSPVASDVSSF

ASALPRKTSLSFSLPQDELEPVKKWRGDLNGLLVAPRNACTVHEDV

RAAAGPKDMEETLTCICODLLHDCVSLQPCWHTFCAACYSGMWRSSLCCTCRCPV

ERICKHLNLNVEVLIQHPDKSRSEEDVQSDARNKITODMLQPKVRSFSEGS

SEDLLESDVDSSESDISQPYVVCRCQPEYRQQAQPHCPAPGEFGAPQALQDAPS

TSVSLTAVDQVYCPQGSQHALCTCCFPMPDRARERODPRAPQCAVCLQPFCHL

YWGCTGCGCLAPFCELNGLKLDGLVNNNSYSELTINKYLATRLGLTKWNLMTES

LVALQGVFLDSLVRTGTVLVCYCGGLRSFELTYQYRQNIASELPVAVTSRDCY

WGRNCRTOVKAHAMFNHCQTRFKN"

ORIGIN

Query Match 79.6%; Score 2133; DB 9; Length 2639;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 529 GATACTCAGGTGAGGTGAGGGCGAGGGCGAGTCCCGGGTCCCTCCGTCGTCGCC 588
Db 481 GATACTCAGGTGAGGTGAGGGCGAGGGCGAGTCCCGGGTCCCTCCGTCGTCGCC 540

Qy 589 GCCACTCAGGTGCTTTGAGGAAACACAGCATCAACATGACGTGACACCTCTTCCC 648
Db 541 GCCACTCAGGTGCTTTGAGGAAACACAGCATCAACATGACGTGACACCTCTTCCC 600

Qy 649 ACAGCTTCGGCCCTCTTCCACGAGGCTTCTCTGAGGCGGAGAGGTTCTTCCAGTTGT 708
Db 601 ACAGCTTCGGCCCTCTTCCACGAGGCTTCTCTGAGGCGGAGAGGTTCTTCCAGTTGT 660

Qy 709 GGGTCTGGGGGTGGTGGCATCTCCCTAAAGAGTGTCTCTGTGGCAAGTATGAA 768
Db 661 GGGTCTGGGGGTGGTGGCATCTCCCTAAAGAGTGTCTCTGTGGCAAGTATGAA 720

Qy 769 GTCTCCAGCTTTTCCCTCAGCTCTCCAGACAGAAAGATGCTCTTTCGTCGTTGAA 828
Db 721 GTCTCCAGCTTTTCCCTCAGCTCTCCAGACAGAAAGATGCTCTTTCGTCGTTGAA 780

Qy 829 CCCAGGATCAGAGATTTGGAGCCCGTGAAGAAATGAGAGGAGATGGGACCTT 888
Db 781 CCCAGGATCAGAGATTTGGAGCCCGTGAAGAAATGAGAGGAGATGGGACCTT 840

Qy 889 GACCTGAAACGGGCGAGTTGTTGGTGCACAAACCGCTAGAAATGCCAACCCCTCCACGAG 948
Db 841 GACCTGAAACGGGCGAGTTGTTGGTGCACAAACCGCTAGAAATGCCAACCCCTCCACGAG 900

Qy 949 GAGCTCAGAGCAGCGCTGGGAGCCAGACAGATGGAGAGCGGTGACATGATATC 1008
Db 901 GAGCTCAGAGCAGCGCTGGGAGCCAGACAGATGGAGAGCGGTGACATGATATC 960

Qy 1009 TGCAGGACCTGTCACAGCTGCTGAGTTTTCGAGCCCTGTCATGCACACAGTTTCTCGCG 1068
Db 961 TGCAGGACCTGTCACAGCTGCTGAGTTTTCGAGCCCTGTCATGCACACAGTTTCTCGCG 1020

Qy 1069 GCTTGTCTCTCGGCTGAGTGGAGCGCTGCTCCCTGTCTTCTACCTGGGCTGTCGGTG 1128
Db 1021 GCTTGTCTCTCGGCTGAGTGGAGCGCTGCTCCCTGTCTTCTACCTGGGCTGTCGGTG 1080

Qy 1129 GAGCGGATCTGTAACCAACATCTCTCAACACCTCTGGAAGCATATCTCATCCAGAT 1188
Db 1081 GAGCGGATCTGTAACCAACATCTCTCAACACCTCTGGAAGCATATCTCATCCAGAT 1140

Qy 1189 CCAGACAAGTCCGAGTGAAGAGATGTCAAGATGATGGATGCCAGGAATAAATCACT 1248
Db 1141 CCAGACAAGTCCGAGTGAAGAGATGTCAAGATGATGGATGCCAGGAATAAATCACT 1200

Qy 1249 CAAGACATCTCAGCCCAAGTCTGAGGCGTCTTTTCTGATGAAGAGGAGTTCAAG 1308
Db 1201 CAAGACATCTCAGCCCAAGTCTGAGGCGTCTTTTCTGATGAAGAGGAGTTCAAG 1260

Qy 1309 GACCTGCTGAGCTGTCAGCGTTGACAGTGGTCTCAGACATTTAGCCAGCCATACGTC 1368
Db 1261 GACCTGCTGAGCTGTCAGCGTTGACAGTGGTCTCAGACATTTAGCCAGCCATACGTC 1320

Qy 1369 GTGTGCGGAGTCTCTGAGTACAGAGGCGGCGGCGGAGCTTCCCTGAGCCAGCA 1428
Db 1321 GTGTGCGGAGTCTCTGAGTACAGAGGCGGCGGCGGAGCTTCCCTGAGCCAGCA 1380

Qy 1429 CCGAGGGGAGCCAGAGGCCCCACAGGCTTGGGGATGACACCTCCACGCTCCGTGAGC 1488
Db 1381 CCGAGGGGAGCCAGAGGCCCCACAGGCTTGGGGATGACACCTCCACGCTCCGTGAGC 1440

Qy 1489 CTGACGACAGCTCAGGATTTAGCTGCTCTGCAAGAGCAAGCCAGCCCTGTGACCC 1548
Db 1441 CTGACGACAGCTCAGGATTTAGCTGCTCTGCAAGAGCAAGCCAGCCCTGTGACCC 1500

Qy	1549	TGCTGCTTCCAGCCCATGCTCCGACCGGAGAGCGGAGCGAGACCCCGCGTGTGCCC	1608
Db	1501	TGCTGCTTCCAGCCCATGCTCCGACCGGAGAGCGGAGCGGAGACCCCGCGTGTGCCC	1560
Qy	1609	CCTCAGCAGTGTGGGCTCTGCTCAGCCCTTCTGCACCTCTACTGTGGGGTGCACCCGG	1668
Db	1561	CCTCAGCAGTGTGGGCTCTGCTCAGCCCTTCTGCACCTCTACTGTGGGGTGCACCCGG	1620
Qy	1669	ACGGCTGTAGGCTGCTGGCCCGCTTTGTGAGCTCAACTGGGTGACAGTGTCTG	1728
Db	1621	ACGGCTGTAGGCTGCTGGCCCGCTTTGTGAGCTCAACTGGGTGACAGTGTCTG	1680
Qy	1729	GACGGCTGCTGAACAACAACAGCTACAGCTCAGACATCCTCAAGAAATTAACCTGGCAACC	1788
Db	1681	GACGGCTGCTGAACAACAACAACAGCTACAGCTCAGACATCCTCAAGAAATTAACCTGGCAACC	1740
Qy	1789	AGAGGTTTGAATCGAATAAACAATGTCGACGAGGCTCTGGCTCTCCACGGGGAGTG	1848
Db	1741	AGAGGTTTGAATCGAATAAACAATGTCGACGAGGCTCTGGCTCTCCACGGGGAGTG	1800
Qy	1849	TTTCTGCTGTCTGATTACAGAGTCACGGAGACACCGTTCTGTGTACTGTGTGGGCTG	1908
Db	1801	TTTCTGCTGTCTGATTACAGAGTCACGGAGACACCGTTCTGTGTACTGTGTGGGCTG	1860
Qy	1909	CGCAGTCTCCGTGAGCTGACCTATCAGTATCGCGAGAACATTCCTGCTCCGAGTGCACA	1968
Db	1861	CGCAGTCTCCGTGAGCTGACCTATCAGTATCGCGAGAACATTCCTGCTCCGAGTGCACA	1920
Qy	1969	GTGGCGGTAAACATTCGGTCTGACTGCTACTGGGGCCGTAACTCGCGCACTCAGGTGAAA	2028
Db	1921	GTGGCGGTAAACATTCGGTCTGACTGCTACTGGGGCCGTAACTCGCGCACTCAGGTGAAA	1980
Qy	2029	GCTCACCACGCGCATGAATTCATATCTGTGAACAGACAAGGTTCAAAAAATAAGCA	2088
Db	1981	GCTCACCACGCGCATGAATTCATATCTGTGAACAGACAAGGTTCAAAAAATAAGCA	2040
Qy	2089	TCCAGAGCCCTGAGCAGCTTTCAGCACTGGAGGTGAAGAGAGCGTGTTTTAAATACA	2148
Db	2041	TCCAGAGCCCTGAGCAGCTTTCAGCACTGGAGGTGAAGAGAGCGTGTTTTAAATACA	2100
Qy	2149	GAGACAAGCAGCTCAAGGTGTTTTTCAGCCGCCCTTGAGGGAAGGACGACGGGTCTCCGA	2208
Db	2101	GAGACAAGCAGCTCAAGGTGTTTTTCAGCCGCCCTTGAGGGAAGGACGACGGGTCTCCGA	2160
Qy	2209	CAGGTGCTCTGGGGTGACTCTCTGTGGAGCTTTTACCCTCTGAGTGAGACCTCCCCA	2268
Db	2161	CAGGTGCTCTGGGGTGACTCTCTGTGGAGCTTTTACCCTCTGAGTGAGACCTCCCCA	2220
Qy	2269	GAGCCCCGGGGCCGACGCCCCCTCTGTGAGCGCTGGCAGGGCTCGTGTGGCAT	2328
Db	2221	GAGCCCCGGGGCCGACGCCCCCTCTGTGAGCGCTGGCAGGGCTCGTGTGGCAT	2280
Qy	2329	CAGCAGCAGAGACGAAGCCTTTCTGTAAACATGGCGCGCTCCGCGCAGAGGGGCGATTTT	2388
Db	2281	CAGCAGCAGAGACGAAGCCTTTCTGTAAACATGGCGCGCTCCGCGCAGAGGGGCGATTTT	2340
Qy	2389	GCTCTTTTGTACATTTCCGAAACTACAGTTAAGCAGAGTCTGTTTTCAGGAAAAAGTT	2448
Db	2341	GCTCTTTTGTACATTTCCGAAACTACAGTTAAGCAGAGTCTGTTTTCAGGAAAAAGTT	2400
Qy	2449	TCAAGCGAAGAAGGCAAGTTTATCAAAAAACATTTTTCAGAGAGGAGGAGCATAAAGTTTA	2508
Db	2401	TCAAGCGAAGAAGGCAAGTTTATCAAAAAACATTTTTCAGAGAGGAGGAGCATAAAGTTTA	2460
Qy	2509	GAGCCTACAGACGTACACAATATCTGCTGGGGAACCAACAGCATTTTATCTATTTT	2568
Db	2461	GAGCCTACAGACGTACACAATATCTGCTGGGGAACCAACAGCATTTTATCTATTTT	2520
Qy	2569	TTTATTTTAAATAGTTTGGTGCTTATCTCTAATAAGATTTAAATGTACAAACTGTAGC	2628
Db	2521	TTTATTTTAAATAGTTTGGTGCTTATCTCTAATAAGATTTAAATGTACAAACTGTAGC	2580

QY	2629	ACAAATAATATATTATTTATTAATTTACAAGTGC	2661
DB	2581	ACAAATAATATATTATTTATTAATTTACAAGTGC	2613

RESULT 5	
LOCUS	BC012072
DEFINITION	Homo sapiens checkpoint with forkhead and ring finger domains, mRNA (CDNA clone MGC:19963 IMAGE:4650348), complete cds.
ACCESSION	BC012072
VERSION	BC012072.1
KEYWORDS	GI:15082329
SOURCE	MGC.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 3189)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Bustow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.P., Jordan,H., Moore,T.J., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marudina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McSwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.C., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A.C., Rodrigues,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butcherfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E., Sengerich,A., Schein,J.E., Jones,S.J. and Marra,M.A., 2000 full-length generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
TITLE	Genet. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 23398257
JOURNAL MEDLINE	2 (bases 1 to 3189)
PUBMED	Strausberg,R.
REFERENCE	Direct Submission
AUTHORS	Submitted (02-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov
REMARK COMMENT	Contact: MGC help desk Email: cgabs@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Boadet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kusche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McEavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Farvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 29 Row: n Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8922674.
Location/Qualifiers
1. -3189

FEATURES

/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="MGC:19963 IMAGE:4650348"
/issue_type="Placenta, Chorioncarcinoma"
/clone_lib="NIH MGC 21"
/lab_host="DH10B-R"
/note="Vector: pOT7"
1. 3189
/gene="CHFR"
/note="synonyms: FLJ10796, RNF116"
/db_xref="LocusID:55743"
/db_xref="MIM:605209"
9. 1967
/codon_start=1
/product="CHFR protein"
/protein_id="AAH12072.1"
/db_xref="GI:15082330"
/db_xref="LocusID:55743"
/translation="MERPEEGKPPPPQWGLRLRLGAEAGEPHVLLKREWTIGRRR
GDLSPFNKLVSGDHCRIVVDEKSGVLTEDTSGTVINKLVKVPKQTCPLQGV
IYLVYRNEPHNVAVLYESLEKGMTOESFDSGAGAGADPRVPPSPATQVCF
PEPQSTSLDLPFASASSTEPSPAGRSSCGSGGGISPKSGSPVASDRVSF
ASALDRKTAFFSLPEQDLELPVKKXRGDGLDLNGQLLVAQPRRNQVTHEDV
RAAGKPKMETLTCIIQDLDHCVSLQPCMETFCAACYSYGMWRSRSLQTCRCVP
ERICNHLINLVEAYLIQHPDKSRSEEDVQMDARNKI TDMLQPKVRSFSDSEGS
SEDLLESDVSESDISQPVVCRQCEYRROAAQPHCPAPEGEPAQALGDAPS
TSVSLTAVQDVPCFGLNLGDKCLDGLVANNVSIEDILKNYLATRLGLTWKMLTES
YWGCTRTCYGLA PFCGLNLGDKCLDGLVANNVSIEDILKNYLATRLGLTWKMLTES
LVALQGVFLSLDVRVGTDLVLCVCGGLRSFRELTYQYRQNI PASELPVAVTSRPDCY
WGRNCRTOVKAHHAWKFNHICEQTRFN"
120. 317
/note="FHA; Region: FHA domain. The FHA
(Forhead-associated) domain is a phosphopeptide binding
motif"
/db_xref="CDD:pfam00498"
879. 1010
/note="HRD1; Region: COG5243, HRD1, HRD ubiquitin ligase
complex, ER membrane component [Posttranslational
modification, protein turnover, chaperones]"
/db_xref="CDD:COG5243"

misc_feature
120. 317
/note="FHA; Region: FHA domain. The FHA
(Forhead-associated) domain is a phosphopeptide binding
motif"
/db_xref="CDD:pfam00498"
879. 1010
/note="HRD1; Region: COG5243, HRD1, HRD ubiquitin ligase
complex, ER membrane component [Posttranslational
modification, protein turnover, chaperones]"
/db_xref="CDD:COG5243"

misc_feature
879. 1010
/note="HRD1; Region: COG5243, HRD1, HRD ubiquitin ligase
complex, ER membrane component [Posttranslational
modification, protein turnover, chaperones]"
/db_xref="CDD:COG5243"

ORIGIN

Query Match 75.8%; Score 2031; DB 9; Length 3189;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2131; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 529 GATACCTCAGGTGACGGTCCAGGGGAGGGGCGGATCCCGGTCCTCCGTCGTCGCC 588
Db 411 GATACCTCAGGTGACGGTCCAGGGGAGGGGCGGATCCCGGTCCTCCGTCGTCGCC 470
Qy 589 GCCACTCAGGTGCTTTGAGGAACCAAGCCATCAACATCGACGTGACCTCTTCCC 648
Db 471 GCCACTCAGGTGCTTTGAGGAACCAAGCCATCAACATCGACGTGACCTCTTCCC 530
Qy 649 ACAGCCTCGGCTCTTCCAGGAGCTTCTCTCCAGGGGAGAGCGTTCTCCAGTTGT 708
Db 531 ACAGCCTCGGCTCTTCCAGGAGCTTCTCTCCAGGGGAGAGCGTTCTCCAGTTGT 590
Qy 709 GGGTCGCGGGTGGTGGCATCTCCCTAAAGAAAGTGGTCCCTCTGTGGCAAGTAA 768
Db 591 GGGTCGCGGGTGGTGGCATCTCCCTAAAGAAAGTGGTCCCTCTGTGGCAAGTAA 650
Qy 769 GTCTCAGCTTTGCCCTCAGCTCTCCAGACAGAAAGCTCGCTCTTTTCGTCGTGAA 828
Db 651 GTCTCAGCTTTGCCCTCAGCTCTCCAGACAGAAAGCTCGCTCTTTTCGTCGTGAA 710
Qy 829 CCCAGGATCAGAGGATTTGAGCCCGTGAAGAAAGAAATAGAGGAGATGGGACCTT 888
Db 711 CCCAGGATCAGAGGATTTGAGCCCGTGAAGAAAGAAATAGAGGAGATGGGACCTT 770
Qy 889 GACCTGAAAGGCGAGTTGTTGTCACACCGCTAGAAATGCCAAACCGTCCACGAG 948

771 GACCTGAAAGGCGAGTTGTTGGTCGCAAAACCGCTAGAAATGCCAAACCGTCCACGAG 830
Qy 949 GACCTGAAAGGCGAGTTGTTGGGAGAACCGACAGATGGAGGAGACCGTGCATCATC 1008
Db 831 GACCTGAAAGGCGAGTTGTTGGGAGAACCGACAGATGGAGGAGACCGTGCATCATC 890
Qy 1009 TGCAGGACCTGCTGCACGACTGCGTAGTTTGGCAGCCCTGCTGATGACACAGTTTCTGCGG 1068
Db 891 TGCAGGACCTGCTGCACGACTGCGTAGTTTGGCAGCCCTGCTGATGACACAGTTTCTGCGG 950
Qy 1069 GCTTGTACTCGGGCTGGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1128
Db 951 GCTTGTACTCGGGCTGGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1010
Qy 1129 GAGCGGATCTGTAATAAACCAATCTCTCAACACCTCTGCTGGAAGCATACCTCATCCAGCAT 1188
Db 1011 GAGCGGATCTGTAATAAACCAATCTCTCAACACCTCTGCTGGAAGCATACCTCATCCAGCAT 1070
Qy 1189 CCAGACAGAGTGCAGTGAAGAGATGTGCAAGATGTGCAAGATGTGCAAGATGTGCAAGATGTG 1248
Db 1071 CCAGACAGAGTGCAGTGAAGAGATGTGCAAGATGTGCAAGATGTGCAAGATGTGCAAGATGTG 1130
Qy 1249 CAAGACATGCTGAGCGCCCAAGAGTCAAGCGGTCTTTTCTGATGAAGAGAGATTTGAGAG 1308
Db 1131 CAAGACATGCTGAGCGCCCAAGAGTCAAGCGGTCTTTTCTGATGAAGAGAGATTTGAGAG 1190
Qy 1309 GACCTGCTGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 1368
Db 1191 GACCTGCTGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 1250
Qy 1369 GTGTGCGGCGAGTCTCTGAGTACAGAGCGAGCGGCGGCGAGCTCCCTCCACTGCCCAGCA 1428
Db 1251 GTGTGCGGCGAGTCTCTGAGTACAGAGCGAGCGGCGGCGAGCTCCCTCCACTGCCCAGCA 1310
Qy 1429 CCCAGGCGGAGCCAGGAGCCCAAGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 1488
Db 1311 CCCAGGCGGAGCCAGGAGCCCAAGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 1370
Qy 1489 CTGACAGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 1548
Db 1371 CTGACAGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 1430
Qy 1549 TGCTGCTTCCAGCGCCATGCGGACCGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGTGTG 1608
Db 1431 TGCTGCTTCCAGCGCCATGCGGACCGGAGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTG 1490
Qy 1609 CTTGACAGTGTGCGGTGCTGCTGAGCGCTTTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1668
Db 1491 CTTGACAGTGTGCGGTGCTGCTGAGCGCTTTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1550
Qy 1669 ACCGGTGTCTACGGCTGCTGCGCCCGCTTTTGTGAGCTCAACCTGGGTGACAGTGTGCTG 1728
Db 1551 ACCGGTGTCTACGGCTGCTGCGCCCGCTTTTGTGAGCTCAACCTGGGTGACAGTGTGCTG 1610
Qy 1729 GAGCGGTGCTGAACCAACAGACTACGACTACGACTACGACTACGACTACGACTACGACTACGACTACG 1788
Db 1611 GAGCGGTGCTGAACCAACAGACTACGACTACGACTACGACTACGACTACGACTACGACTACGACTACG 1670
Qy 1789 AGAGTTTGAATGGAATAAATCATGTTGACCGAGAGCTCTGCTGCTCTCCAGCGGGAGTG 1848
Db 1671 AGAGTTTGAATGGAATAAATCATGTTGACCGAGAGCTCTGCTGCTCTCCAGCGGGAGTG 1730
Qy 1849 TTTCTG 1908
Db 1731 TTTCTG 1790
Qy 1909 CGCAGCTTCCGCTG 1968
Db 1791 CGCAGCTTCCGCTG 1850
Qy 1969 GTGCGCTTAACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2028
Db 1851 GTGCGCTTAACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1910

QY 2629 ACAATAATATATATTATTAATTTACAAATTGAC 2661
 |||||
 DB 2509 ACAATAATATATATTATTAATTTACAAATTGAC 2541
 |||||

RESULT 7
 HSM802304
 LOCUS
 DEFINITION Homo sapiens mRNA; cDNA DKFPz434N2420 (from clone DKFZp434N2420); linear PRI 18-FEB-2000
 Partial cds.
 ACCESSION AL137561
 VERSION AL137561.1 GI:6808265
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2297)
 AUTHORS Ottenwälder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp434N2420) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cdna/. Location/Qualifiers
 1..2297
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp434N2420"
 /tissue_type="testis"
 /clone_join="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1..2297
 /gene="DKFZp434N2420"
 <1..921
 /gene="DKFZp434N2420"
 /note="unknown"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CA570812.1"
 /db_xref="GI:6808266"
 /db_xref="SPTREMBL:Q9NTJ32"
 /translation="VSAYLIHQDPKDSSEEDVQSDPARNKITQDMLQPKVRFRSFDEE GSSEDLLSDVDVSCPSQPVRCQCPEYRQAQPHPPEGEPAQALGAD PNSTSVLTAVQDVVCPLOQSHALCTCCFOPMPDRARERQDPRAVQQCAVCLOPPC HLWGTGTCGYCLAPFCENLNGDKLGDVLANNVSYESDLKNYLATRGLTWKNNLT ESLVAGRGVFLLSDRYVTGCTVLCVCGGLSRFELTYQYQNIPASELPFAVTSRSD CYWRNCRTGVKAHKAMFKFNHCQTRFXN"
 polyA_signal 2227..2232
 /gene="DKFZp434N2420"
 polyA_site 2251
 /gene="DKFZp434N2420"

ORIGIN

Query Match	55.9%; Score 1497; DB 9; Length 2297;
Best Local Similarity	100.0%; Prematch.No.: 0;
Matches 1497; Conservative	0; Mismatches 0; Indels 0; Gaps 0;

QY 1165 GTGGAAGCACCTCATCCAGCATCCAGACAAGAGTGCAGTGCAAGATGTCAAAGT 1224
 DB 1 GTGGAGCATACCTCATCCAGCATCCAGACAAGAGTGCAGTGCAAGATGTCAAAGT 60

Db 1141 GCTGGCAGGCGCTGTTGGTGGTCATCAGCAGCAGAGCAAGCCCTTTCTGTAAATGCGGC 1200
Qy 2365 CGTCCCGCCGAGAGGCGCAGTTTGTCTCTTTTGTACATTTTTCGAAACTACAGTTAAAGC 2424
Db 1201 CGTCCCGCCGAGAGGCGCAGTTTGTCTCTTTTGTACATTTTTCGAAACTACAGTTAAAGC 1260
Qy 2425 AGAAGTCTGTTTTTCAGGAAAGTTTCAAGGAGAGGAGGAGGAGTTTATCAAAAACATTTT 2484
Db 1261 AGAAGTCTGTTTTTCAGGAAAGTTTCAAGGAGAGGAGGAGGAGTTTATCAAAAACATTTT 1320
Qy 2485 TCAGGAGAGGAGCATAAGTTTACAGCCTACAGCAGTACACAAATATCTCTGCTGCTGGG 2544
Db 1321 TCAGGAGAGGAGCATAAGTTTACAGCCTACAGCAGTACACAAATATCTCTGCTGCTGGG 1380
Qy 2545 AAAACCCAGCAGATTTTATCTATTTTATTTTAAATAGTTTGGTGTATCTTTCTAATAA 2604
Db 1381 AAAACCCAGCAGATTTTATCTATTTTATTTTAAATAGTTTGGTGTATCTTTCTAATAA 1440
Qy 2605 GATTAAATGTCACAACTGTAGCACAATAATATTAATTTTAAATTTACAAATTTGAC 2661
Db 1441 GATTAAATGTCACAACTGTAGCACAATAATATTAATTTTAAATTTTACAAATTTGAC 1497

RESULT 8
AX877222
LOCUS AX877222 3138 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 12127 from Patent EP1074617.
ACCESSION AX877222
VERSION AX877222.1 GI:40031958
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primers for synthesising full-length cDNA and their use
Patent: EP 1074617-A 12127 07-FEB-2001;
Research Association for Biotechnology (JP)

FEATURES
source
1..3138
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
65..1936
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB89626.1"
/translation="MERPEKGQPPPPQPPWGRLLRLGAEAGEPHVLLKREWTIGRRV
CGLSPSPSKLVSGDHCHIVVDEKSGOVLTEDTSTGTVINKLVKKQTCPLQTDV
IYLVRKNEPHNVAYLSELSKQMTQESFEMVPCVQAQGLKLGSSDPLEPKK
SIVITGSGGIGSPGSPSVASDEVSPFALPDKTAFSPFSLPEQDQDLEPKK
MRGDDLLNGOLLVAQPRNAQTVHEDVRAAGKPKDMEETLILICODLLHDCVSL
QPCMTFCAACYSNMERSLCTCRCEPERIKNNHILNNLVEAYLIORPKSRSEED
VQSDARNKITQDMLQPKVRSFSDDESSDLELSDVDESSDIQFVVCRCQPE
YRQAAQPHCPAPEGPAQALDAPPTSVSLTAVQDYVCPQGGSHALCTCCFP
NNRREAREQDPVPAQCAVCLQPFCHLYGCTRTGCGYGLAFCELNLDGKLDGV
LNNRSEEDILKLVATRLGLTKNMLTBSLVALQRFVLLSDYRVGTGLVLCVCCGLR
SPRELYTQYQNIIPASELPVATSRPDCYWGNCRTQVKAHAKFNHICEOTRPN"

CDS
1..3138
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
65..1936
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB89626.1"
/translation="MERPEKGQPPPPQPPWGRLLRLGAEAGEPHVLLKREWTIGRRV
CGLSPSPSKLVSGDHCHIVVDEKSGOVLTEDTSTGTVINKLVKKQTCPLQTDV
IYLVRKNEPHNVAYLSELSKQMTQESFEMVPCVQAQGLKLGSSDPLEPKK
SIVITGSGGIGSPGSPSVASDEVSPFALPDKTAFSPFSLPEQDQDLEPKK
MRGDDLLNGOLLVAQPRNAQTVHEDVRAAGKPKDMEETLILICODLLHDCVSL
QPCMTFCAACYSNMERSLCTCRCEPERIKNNHILNNLVEAYLIORPKSRSEED
VQSDARNKITQDMLQPKVRSFSDDESSDLELSDVDESSDIQFVVCRCQPE
YRQAAQPHCPAPEGPAQALDAPPTSVSLTAVQDYVCPQGGSHALCTCCFP
NNRREAREQDPVPAQCAVCLQPFCHLYGCTRTGCGYGLAFCELNLDGKLDGV
LNNRSEEDILKLVATRLGLTKNMLTBSLVALQRFVLLSDYRVGTGLVLCVCCGLR
SPRELYTQYQNIIPASELPVATSRPDCYWGNCRTQVKAHAKFNHICEOTRPN"

ORIGIN
Query Match 54.7%; Score 1465; DB 6; Length 3138;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1945; Conservative 0; Mismatches 6; Indels 2; Gaps 1;
Qy 709 GGGCTGGGGTGGTGGCATCTCCCTAAAGCAAGTGTCCCTCTGTGGCAAGTGAATGAA 768
Db 560 GGGCTGGGGTGGTGGCATCTCCCTAAAGCAAGTGTCCCTCTGTGGCAAGTGAATGAA 619
Qy 769 GTCTCCAGCTTTGGCTCAGCTCTCCAGACAGAAAGATGGCTCTTTTGTGTTGGA 828

Db 61 ATGGATGCCAGGATTAATCACTCAAGACATGCTGCAGCCCAAGTCAAGCGGCTTTT 120
Qy 1285 TCTGATGAAGAGGAGTTCAAGGACCTCTGAGACCTCTCAGACCTTGAACAGTGAAGTCC 1344
Db 121 TCTGATGAAGAGGAGTTCAAGGACCTCTGAGACCTCTCAGACCTTGAACAGTGAAGTCC 180
Qy 1345 TCAGACATTAGCAGCAGCAGTCTGCTGCGGAGTCTCTGAGTACAGAAAGGAGGCG 1404
Db 181 TCAGACATTAGCAGCAGCAGTCTGCTGCGGAGTCTCTGAGTACAGAAAGGAGGCG 240
Qy 1405 GCGCAGCCTCCCACTGCCCCAGCAGCCGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1464
Db 241 GCGCAGCCTCCCACTGCCCCAGCAGCCGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Qy 1465 GATGACCTCTCAGCTCCGTCAGCTGACAGCAGTCCAGGATTAAGTGTGCTGCTCTG 1524
Db 301 GATGACCTCTCAGCTCCGTCAGCTGACAGCAGTCCAGGATTAAGTGTGCTGCTCTG 360
Qy 1525 CAAGGAAGCAGCAGCCTCTGCTGCTGCTTCCAGCCCAATGCCCCAGGAGGAGGAGGAG 1584
Db 361 CAAGGAAGCAGCAGCCTCTGCTGCTGCTTCCAGCCCAATGCCCCAGGAGGAGGAGGAG 420
Qy 1585 GCGCAGCAGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
Db 421 GCGCAGCAGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 1645 CACTGTACTGCGGCTGACCCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704
Db 481 CACTGTACTGCGGCTGACCCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 1705 CTCACCTGGGTGACAGTGTCTGAGCGGCTGCTGAGCAACCAACAGCTACAGTCAAGC 1764
Db 541 CTCACCTGGGTGACAGTGTCTGAGCGGCTGCTGAGCAACCAACAGCTACAGTCAAGC 600
Qy 1765 ATCTGAGAAATACCTGCAACAGAGGTTGACATGGAACCAACAGCTGTTGACCGAGAGC 1824
Db 601 ATCTGAGAAATACCTGCAACAGAGGTTGACATGGAACCAACAGCTGTTGACCGAGAGC 660
Qy 1825 CTCCTGCTCTCCAGCGGAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1884
Db 661 CTCCTGCTCTCCAGCGGAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 1885 GTTCTGTGTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
Db 721 GTTCTGTGTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 1945 AACATTCTGCTTCCGAGTTGCCAGTGGCGGTAAACATCCCGTCTGCTGCTGCTGCTGCTG 2004
Db 781 AACATTCTGCTTCCGAGTTGCCAGTGGCGGTAAACATCCCGTCTGCTGCTGCTGCTGCTG 840
Qy 2005 CGTAACTCCCGACTCAGGTGAAGCTCAACAGCCATGAATTCATATCTGTGAA 2064
Db 841 CGTAACTCCCGACTCAGGTGAAGCTCAACAGCCATGAATTCATATCTGTGAA 900
Qy 2065 CAGCAAGGTTCAAAACTAAGCATCCAGAGGCGCTGAGCAGCTTTTCAGCACTGGAGGTG 2124
Db 901 CAGCAAGGTTCAAAACTAAGCATCCAGAGGCGCTGAGCAGCTTTTCAGCACTGGAGGTG 960
Qy 2125 AAGAGAGCGTGTGTTTAAATACAGAGCAAGCAGTCAAGGTTGTTTTCACAGCCCGCTG 2184
Db 961 AAGAGAGCGTGTGTTTAAATACAGAGCAAGCAGTCAAGGTTGTTTTCACAGCCCGCTG 1020
Qy 2185 AGGAAGGAGCGAGGTTCTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2244
Db 1021 AGGAAGGAGCGAGGTTCTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Qy 2245 ACCCTCTGAGTGAGACCTCTCCAGAGCGCGCGGCGGAGCGCGCGCTCTCTGCTGAGC 2304
Db 1081 ACCCTCTGAGTGAGACCTCTCCAGAGCGCGCGGCGGAGCGCGCGCTCTCTGCTGAGC 1140
Qy 2305 GCTGGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2364

Db 620 GTCTCAGCTTTGCTCAGCTCTCCAGACAGAAAGACTGGCTCTCTTTTCGCTGTTGAA 679
Qy 829 CCCAGGATCAGGAGATTTGAGCCCGTGAAGAGAAATGAGAGGAGATGGGACCTT 888
Db 680 CCCAGGATCAGGAGATTTGAGCCCGGAGAGAAATGAGAGGAGATGGGACCTT 739
Qy 889 GACCTGAACGGGACGTTGTTGTCGACAAACCGCTAGAAATGCCAAACCGTCCACGAG 948
Db 740 GACCTGAACGGGACGTTGTTGTCGACAAACCGCTAGAAATGCCAAACCGTCCACGAG 799
Qy 949 GAGCTCAGAGACGGCTGGGAGCCAGACAGATGAGAGAGCCTGACATGCATCATC 1008
Db 800 GAGCTCAGAGACGGCTGGGAGCCAGACAGATGAGAGAGCCTGACATGCATCATC 859
Qy 1009 TGCAGGACCTGTGCAACGCTGCGTGAAGTTTGCAGCCCTGATGCACACAGCTTCTGCGCG 1068
Db 860 TGCAGGACCTGTGCAACGCTGCGTGAAGTTTGCAGCCCTGATGCACACAGCTTCTGCGCG 919
Qy 1069 GCTTGCTACTCGGGCTGAGTGAAGCGCTCGTCCCTGTGTCCTACTGCGCTGTCGCCGTG 1128
Db 920 GCTTGCTACTCGGGCTGAGTGAAGCGCTCGTCCCTGTGTCCTACTGCGCTGTCGCCGTG 979
Qy 1129 GAGCGGATCTGTAAACCAACCATCTCTCAACACCTCTGGAAGCATCTCATCCAGCAT 1188
Db 980 GAGCGGATCTGTAAACCAACCATCTCTCAACACCTCTGGAAGCATCTCATCCAGCAT 1039
Qy 1189 CCAGACAGAGTCGAGTGAAGAAGATGTGAAAGTATGATGCCAGGAATAAATCACT 1248
Db 1040 CCAGACAGAGTCGAGTGAAGAAGATGTGAAAGTATGATGCCAGGAATAAATCACT 1099
Qy 1249 CAAGACATCTGAGCCCAAGTCAAGCGCTCTTTCTGATGAAGAGGAGTTCAAG 1308
Db 1100 CAAGACATCTGAGCCCAAGTCAAGCGCTCTTTCTGATGAAGAGGAGTTCAAG 1159
Qy 1309 GACCTCTCGAGCTGTGACAGCTGTGACAGTGTCTCAGACATTTAGCCAGCCATACCTC 1368
Db 1160 GACCTCTCGAGCTGTGACAGCTGTGACAGTGTCTCAGACATTTAGCCAGCCATACCTC 1219
Qy 1369 GTGTGCGGAGTGTCTGATGACAGAGCGGCGGCGACCTCTCCACTGCCCCAGCA 1428
Db 1220 GTGTGCGGAGTGTCTGATGACAGAGCGGCGGCGACCTCTCCACTGCCCCAGCA 1279
Qy 1429 CCCAGGGCGAGCAGGAGCCCCACAGGCCCTGGGGGATGACCTCTCAAGCTCCGTCAGC 1488
Db 1280 CCCAGGGCGAGCAGGAGCCCCACAGGCCCTGGGGGATGACCTCTCAAGCTCCGTCAGC 1339
Qy 1489 CTGACGACAGCTCCAGATTAAGTGTGCTCTGCAAGAGAGCGACGCCCTGTGACCC 1548
Db 1340 CTGACGACAGCTCCAGATTAAGTGTGCTCTGCAAGAGAGCGACGCCCTGTGACCC 1399
Qy 1549 TGCTGCTTCCAGCCATGCTGAGCGAGCGGAGCGAGCGAGCGGAGCGGAGCGGAGCGGAG 1608
Db 1400 TGCTGCTTCCAGCCATGCTGAGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1459
Qy 1609 CCTCAGAGTGTGCGTCTGCTGAGCCTTTCTGCACTGTGAGCGGCTGCAACCGG 1668
Db 1460 CCTCAGAGTGTGCGTCTGCTGAGCCTTTCTGCACTGTGAGCGGCTGCAACCGG 1519
Qy 1669 ACCGCTGTACGGCTGCTGCGCCCGTGTGAGCTCAACTGAGTGTGAGTGTGAGTGTGAG 1728
Db 1520 ACCGCTGTACGGCTGCTGCGCCCGTGTGAGCTCAACTGAGTGTGAGTGTGAGTGTGAG 1579
Qy 1729 GAGCGGCTGTGAACCAACACAGTACAGTACAGATCTGAGAGATTTACTTGCACACC 1788
Db 1580 GAGCGGCTGTGAACCAACACAGTACAGTACAGATCTGAGAGATTTACTTGCACACC 1639
Qy 1789 AGAGTTTGAATGAGAAATGTTGACCGAGAGCTGCTGCTCTCCAGCGGGAGTG 1848
Db 1640 AGAGTTTGAATGAGAAATGTTGACCGAGAGCTGCTGCTCTCCAGCGGGAGTG 1699
Qy 1849 TTTCTGCTGTGATTAACAGTCAAGGAGACCGGTTCTGTGTTACTGCTGTGCGCTG 1908
Db 1700 TTTCTGCTGTGATTAACAGTCAAGGAGACCGGTTCTGTGTTACTGCTGTGCGCTG 1759

Qy 1909 CGCAGCTTCCGTGAGCTGACCTATCAGTATCGGCAACAATTCCTCTCTCCGAGTTGCCA 1968
Db 1760 CGCAGCTTCCGTGAGCTGACCTATCAGTATCAGCAACAATTCCTCTCTCCGAGTTGCCA 1819
Qy 1969 GTGCGCTTAACATCCCTGCTGACTCTGACTCTGGGCGGTAACCTGCGGCACTCAGGTGAAA 2028
Db 1820 GTGCGCGTAAATCCCTGCTGACTCTGACTCTGGGCGGTAACCTGCGGCACTCAGGTGAAA 1879
Qy 2029 GCTCACCACGCCATGAATTCATATCTGTGAACAGCAAGGTTCAAAAATCAAGCA 2088
Db 1880 GCTCACCACGCCATGAATTCATATCTGTGAACAGCAAGGTTCAAAAATCAAGCA 1939
Qy 2089 TCAGAGGCGCTCAGAGAGCTTTTCAAGCACTGAGGTGAAGAGCGTGTGTTTAAATACA 2148
Db 1940 TCAGAGGCGCTCAGAGAGCTTTTCAAGCACTGAGGTGAAGAGCGTGTGTTTAAATACA 1999
Qy 2149 GAGCAAGCAGCTCAAGGTGTTTTCAGCGCCCTCAGGGAAGGAGCGCAGGCTCTCCGA 2208
Db 2000 GAGCAAGCAGCTCAAGGTGTTTTCAGCGCCCTCAGGGAAGGAGCGCAGGCTCTCCGA 2059
Qy 2209 CAGGTGCTCTGGGCTGACTCTTCTGTGAGCTTTTACCCCTGTGAGTGAAGCCCTCCCA 2268
Db 2060 CAGGTGCTCTGGGCTGACTCTTCTGTGAGC--TTTACCCCTGTGAGTGAAGCCCTCCCA 2117
Qy 2269 GAGCCCGGGGCGGAGCGCCCTCTCTGTCAGCGCTGGGCGAGGCTCGTGTGSCAT 2328
Db 2118 GAGCCCGGGGCGGAGCGCCCTCTCTGTCAGCGCTGGGCGAGGCTCGTGTGSCAT 2177
Qy 2329 CAGCAGCAGAGACGAAGCTTTCTGTAACATCGGCGCTCCCGCGAGAGGCGAGTTT 2388
Db 2178 CAGCAGCAGAGACGAAGCTTTCTGTAACATCGGCGCTCCCGCGAGAGGCGAGTTT 2237
Qy 2389 GCTCTTTTGTACATTTTTCGAAACTCAGTTAAAGCAGAGTCTGTTTTCAGGAAGAGTT 2448
Db 2238 GCTCTTTTGTACATTTTTCGAAACTCAGTTAAAGCAGAGTCTGTTTTCAGGAAGAGTT 2297
Qy 2449 TCAGGAGAGAGGCGAGTTTATCAAAAATTTTTCAGGAGAGGAGGAGCATAGTTTA 2508
Db 2298 TCAGGAGAGAGGCGAGTTTATCAAAAATTTTTCAGGAGAGGAGGAGCATAGTTTA 2357
Qy 2509 CAGCCTACAGACGATACAAATATCTGCTGCTGGGAAACCAACAGCAATTTTATCTAT 2568
Db 2358 CAGCCTACAGACGATACAAATATCTGCTGCTGGGAAACCAACAGCAATTTTATCTAT 2417
Qy 2569 TTTATTTTAAATAGTTTGGTGTATCTTCTTAATAGATTTAAATGTCAACAACCTGTAGC 2628
Db 2418 TTTATTTTAAATAGTTTGGTGTATCTTCTTAATAGATTTAAATGTCAACAACCTGTAGC 2477
Qy 2629 ACAATTAATAATTTTAAATTTTACAAATGAC 2661
Db 2478 ACAATTAATAATTTTAAATTTTACAAATGAC 2510

RESULT 9
BD156548
LOCUS
DEFINITION
ACCESSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD156548 3138 bp DNA linear PAT 17-JAN-2003
Primer for synthesizing full-length cDNA and use thereof.
BD156548
ACCESSION BD156548
KEYWORDS BD156548.1 GI:27862306
JP 2002191363-A/11391.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3138)
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 11391 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/11391

PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO.
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUO OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10,
 PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 (65). (1933).

FT CDS Location/Qualifiers
 source
 1. 3138
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 54.7%; Score 1465; DB 6; Length 3138;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1945; Conservative 0; Mismatches 6; Indels 2; Gaps 1;
 709 GGGTCTGGGGTGGTGGCATCTCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTGAATGAA 768
 Db GGGTCTGGGGTGGTGGCATCTCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTGAATGAA 519
 769 GTCTCCAGCTTTGCCCTCAGCTCTCCAGACAGAAAGACTGCGTCTCTTTTCGTGGAA 828
 Db GTCTCCAGCTTTGCCCTCAGCTCTCCAGACAGAAAGACTGCGTCTCTTTTCGTGGAA 679
 829 CCCAGGATCAGAGAGATTGGAGCCCGTGAGAGAAATAGAGAGATGGGACCTT 888
 Db CCCAGGATCAGAGAGATTGGAGCCCGTGAGAGAAATAGAGAGATGGGACCTT 739
 680 CCCAGGATCAGAGAGATTGGAGCCCGTGAGAGAAATAGAGAGATGGGACCTT 739
 889 GACCTGAACCGGCGAGTTGGTGGCAGCAACCGGTAGAAATGCCCAACCGTCCACGAG 948
 Db GACCTGAACCGGCGAGTTGGTGGCAGCAACCGGTAGAAATGCCCAACCGTCCACGAG 799
 949 GACGTGAGAGACGGGTGGGAAGCAGACCAAGATGGAGAGACGCTGACATGCATCATC 1008
 Db GACGTGAGAGACGGGTGGGAAGCAGACCAAGATGGAGAGACGCTGACATGCATCATC 859
 1009 TGCAGAGACCTGTGCACGACTCGGTGAGTTTGCAGCCCTGATGACACAGCTTCTGCGG 1068
 Db TGCAGAGACCTGTGCACGACTCGGTGAGTTTGCAGCCCTGATGACACAGCTTCTGCGG 919
 860 TGCAGAGACCTGTGCACGACTCGGTGAGTTTGCAGCCCTGATGACACAGCTTCTGCGG 919
 1069 GCTTGTCTACTCGGGCTGGATGGAGCGCTGCTCCCTGTGCTTACTGCTGCTGCTCCGCTG 1128
 Db GCTTGTCTACTCGGGCTGGATGGAGCGCTGCTCCCTGTGCTTACTGCTGCTGCTCCGCTG 979
 920 GCTTGTCTACTCGGGCTGGATGGAGCGCTGCTCCCTGTGCTTACTGCTGCTGCTCCGCTG 979
 1129 GAGGGATCTGTAAGAACCAATCTCAACCACTCGTGGAGCATACCTCATCCAGCAT 1188
 Db GAGGGATCTGTAAGAACCAATCTCAACCACTCGTGGAGCATACCTCATCCAGCAT 1039
 980 GAGCGGATCTGTAAGAACCAATCTCAACCACTCGTGGAGCATACCTCATCCAGCAT 1039
 1189 CCAGACAAGAGTCCGAGTGAAGAGATGTGCAAGTATGGATGCCAGGAATAAATCACT 1248
 Db CCAGACAAGAGTCCGAGTGAAGAGATGTGCAAGTATGGATGCCAGGAATAAATCACT 1099
 1249 CAAGACATGCTGAGCCCAAGTCAGCGGTCTTTCTGATGAAGAGGAGTTCAGAG 1308
 Db CAAGACATGCTGAGCCCAAGTCAGCGGTCTTTCTGATGAAGAGGAGTTCAGAG 1159
 1100 CAAGACATGCTGAGCCCAAGTCAGCGGTCTTTCTGATGAAGAGGAGTTCAGAG 1159
 1309 GACCTGCTGGAGTGTGACAGCTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 1368
 Db GACCTGCTGGAGTGTGACAGCTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 1219
 1369 GTGTGCGGAGTGTCTGTGATGACAGAGCAGCGCGGAGCTCCCACTGCCAGCA 1428
 Db GTGTGCGGAGTGTCTGTGATGACAGAGCAGCGCGGAGCTCCCACTGCCAGCA 1279
 1220 GTGTGCGGAGTGTCTGTGATGACAGAGCAGCGCGGAGCTCCCACTGCCAGCA 1279
 1429 CCCGAGGCGGAGCCAGGAGCCCAACAGGCGCTTGGGGGATGACCCCTCCAGCTCCGTCAGC 1488

Db CCCGAGGCGGAGCCAGGAGCCCAACAGGCGCTTGGGGGATGACCCCTCCAGCTCGTCAGC 1339
 QY CTGACGACAGCAGTCCAGGATTTACGTGTGCCCTCTGCAAGGAAGCCACGCGCTGTGACCC 1548
 Db CTGACGACAGCAGTCCAGGATTTACGTGTGCCCTCTGCAAGGAAGCCACGCGCTGTGACCC 1399
 1340 CTGACGACAGCAGTCCAGGATTTACGTGTGCCCTCTGCAAGGAAGCCACGCGCTGTGACCC 1399
 1549 TGCTGCTTCCAGCCATGCGCCGACCGAGAGCGGAGCGCGAGCAGGAGCCCGCTGTGCGC 1608
 1400 TGCTGCTTCCAGCCATGCGCCGACCGAGAGCGGAGCGCGAGCAGGAGCCCGCTGTGCGC 1459
 1609 CCTCAGCAGTGTGGGTCTGCTGCTCAGCCTTTCTGCCCACTGTACTGGGGGTGACCCCGG 1668
 1460 CCTCAGCAGTGTGGGTCTGCTGCTCAGCCTTTCTGCCCACTGTACTGGGGGTGACCCCGG 1519
 1669 ACCGGCTGCTACGGCTGCTGCTGCGCCCTTTGTGAGTCTCAACCTGGGTGACAAAGTGTCTG 1728
 1520 ACCGGCTGCTACGGCTGCTGCTGCGCCCTTTGTGAGTCTCAACCTGGGTGACAAAGTGTCTG 1579
 1729 GACGGCTGCTGACAAACACAGCTACAGTCTCAGACATCTGAGAAATTAACCTGGCAACC 1788
 1580 GACGGCTGCTGACAAACACAGCTACAGTCTCAGACATCTGAGAAATTAACCTGGCAACC 1639
 1789 AGAGGTTTGACATGGAAGAAACATGTTGACCGAGAGCCTCGTGGCTCTCCAGCGGGAGTG 1848
 1640 AGAGGTTTGACATGGAAGAAACATGTTGACCGAGAGCCTCGTGGCTCTCCAGCGGGAGTG 1699
 1849 TTTCTGCTGTGATTAACAGAGTCAAGGAGACACCGTCTGCTGTTACTGCTGTTGGCTG 1908
 1700 TTTCTGCTGTGATTAACAGAGTCAAGGAGACACCGTCTGCTGTTACTGCTGTTGGCTG 1759
 1909 CGCAGCTTCCGTGAGCTGACCTATCAGTATCGGAGAACATTTCTGCTTCCGAGTTGCCA 1968
 1760 CGCAGCTTCCGTGAGCTGACCTATCAGTATCAGCAGAACATTTCTGCTTCCGAGTTGCCA 1819
 1969 GTGGCGGTAAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2028
 1820 GTGGCGGTAAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1879
 2029 GCTCACCAGCCATGAATTAATCATATCTGTAACAGCAAGGTTCAAAATCTAAGCA 2088
 1880 GCTCACCAGCCATGAATTAATCATATCTGTAACAGCAAGGTTCAAAATCTAAGCA 1939
 2089 TCCAGAGCCCTGAGAGCTTTTTCAGCATGAGAGTGAAGAGAGCTGTTTTTAAATACA 2148
 1940 TCCAGAGCCCTGAGAGCTTTTTCAGCATGAGAGTGAAGAGAGCTGTTTTTAAATACA 1999
 2149 GAGCAAGCAGCTCAAGGTGTTTTTTCAGAGCCCTGAGGGAAGGAGCGAGGCTCTCCGA 2208
 2000 GAGCAAGCAGCTCAAGGTGTTTTTTCAGAGCCCTGAGGGAAGGAGCGAGGCTCTCCGA 2059
 2209 CAGGTGCTCTGGGGTGAATCTTCTGAGAGCTTTTACCTCTGAGTGAGACCTTCCCA 2268
 2060 CAGGTGCTCTGGGGTGAATCTTCTGAGAGC - TTTTACCTCTGAGTGAGACCTTCCCA 2117
 2269 GAGCCCGGGGGCGCAGCCCGCTCTGAGAGCTGAGCGCTGAGGCTCGTGTGGCAT 2328
 2118 GAGCCCGGGGGCGCAGCCCGCTCTGAGAGCTGAGCGCTGAGGCTCGTGTGGCAT 2177
 2329 CAGCAGCAGAGAGCAAGGCTTTTCTGTAACATCGCGCGCTCCCGCGAGAGGGGCGAGTTTT 2388
 2178 CAGCAGCAGAGAGCAAGGCTTTTCTGTAACATGCGCGCTCTGCGAGAGGGGCGAGTTTT 2237
 2389 GCTCTTTTGTACATTTTCCGAACTACAGTTAAAGCAGAGCTGTTTTTCAAGGAAAGTT 2448
 2238 GCTCTTTTGTACATTTTCCGAACTACAGTTAAAGCAGAGCTGTTTTTCAAGGAAAGTT 2297
 2449 TCAAGGGAAGAGGGCAAGTTTATCAAAACATTTTTCAGGAGAGGGGAGCATTAAGTTTA 2508
 2298 TCAAGGGAAGAGGGCAAGTTTATCAAAACATTTTTCAGGAGAGGGGAGCATTAAGTTTA 2357
 2509 CAGCCTACAGAGCGTACAAATATCTGCTGCTGGGAAACCAAGCATTTTACTATT 2568

1580 GACGGCTGCTGAACAAACACAGCTACGAGTACAGATCCTGAAGAAATACCTGGCAACC 1639
1789 AGAGGTTTTCACATGGAAAAACATGTTGACCGAGAGCCTCGTGGCTCTCCAGCGGGAGTG 1848
1640 AGAGGTTTTCACATGGAAAAACATGTTGACCGAGAGCCTCGTGGCTCTCCAGCGGGAGTG 1699
1849 TTTCTGCTGTCTGATTTACAGAGTACCGGAGACACCGTTCCTGTGTTACTGCTGTGGCCTG 1908
1700 TTTCTGCTGTCTGATTTACAGAGTACCGGAGACACCGTTCCTGTGTTACTGCTGTGGCCTG 1759
1909 CGCAGCTTCCGTTGAGCTGACCTATCAGTATCGCAGAACATTCCTGCTTCCAGTTGGCA 1968
1760 CGCAGCTTCCGTTGAGCTGACCTATCAGTATCAGCAGAACATTCCTGCTTCCAGTTGGCA 1819
1969 GTGGCGTAACTATCCCGTCTGACTGCTACTGCGGCCGCTAACTGCGCAGCTCAGTGAAA 2028
1820 GTGGCGTAACTATCCCGTCTGACTGCTACTGCGGCCGCTAACTGCGCAGCTCAGTGAAA 1879
2029 GCTCACCGCCATGAATTCATATCATATCTGTGAACACAGAGTTCAAAACCTAAGCA 2088
1880 GCTCACCGCCATGAATTCATATCATATCTGTGAACACAGAGTTCAAAACCTAAGCA 1939
2089 TCCAGAGGCGCTCAGCAGCTTTTCCAGCACTGAGAGTGAAGAGCGTGTGTTTTAAATACA 2148
1940 TCCAGAGGCGCTCAGCAGCTTTTCCAGCACTGAGAGTGAAGAGCGTGTGTTTTAAATACA 1999
2149 GAGACAGACGCTCAGAGTGTGTTTCCAGCGCCCTGAGGAGGAGGAGCGAGGTTCCGA 2208
2000 GAGGCAAGCAGCTCAAGGTGTTTTTCCAGCGCCCTGAGGAGGAGGAGCGAGGTTCCGA 2059
2209 CAGGTGCTCTGGGGTCACTCTTCTGTGGAGCTTTTACCCCTCTGAGTGAGACCCCTCCCA 2268
2060 CAGGTGCTCTGGGGTCACTCTTCTGTGGAGCTTTTACCCCTCTGAGTGAGACCCCTCCCA 2117
2269 GAGCCCCGGGGCCGAGCGCCCTCTGTGGTGGAGCGCTGGGAGGCGTCTGTGGTGGCAT 2328
2118 GAGCCCCGGGGCCGAGCGCCCTCTGTGGTGGAGCGCTGGGAGGCGTCTGTGGTGGCAT 2177
2329 CAGCAGCAGAGCAGAGCTTCTGTGAACATCGCGCCCTCCGCGAGAGGCGAGTGT 2388
2178 CAGCAGCAGAGCAGAGCTTCTGTGAACATCGCGCCCTCCGCGAGAGGCGAGTGT 2237
2389 GCTCTTTTGTACATTTTCGAAACTACAGTTAAAGCAGAGTCTGTGTTTCAGGAAAGTT 2448
2238 GCTCTTTTGTACATTTTCGAAACTACAGTTAAAGCAGAGTCTGTGTTTCAGGAAAGTT 2297
2449 TCAAGGAGAGAGGCGAAGTTTATCAAAACATTTGTTTTCAGGAGAGGAGCATAGTTTA 2508
2298 TCAAGGAGAGAGGCGAAGTTTATCAAAACATTTGTTTTCAGGAGAGGAGCATAGTTTA 2357
2509 CAGCCTACAGAGCGTACAAATATCTGCTGCTGGGAAACCAACAGCATTTTATCTATT 2568
2358 CAGCCTACAGAGCGTACAAATATCTGCTGCTGGGAAACCAACAGCATTTTATCTATT 2417
2569 TTTATTTTAAAGTTTGGTGTATCTTCTTAATAGATTTAAATGTCACAACTGTAGC 2628
2418 TTTATTTTAAAGTTTGGTGTATCTTCTTAATAGATTTAAATGTCACAACTGTAGC 2477
2629 ACAATTAATATATTTATATTTTACAAATTTGAC 2661
2478 ACAATTAATATATTTTATATTTTACAAATTTGAC 2510

RESULT 11
AX834949
LOCUS AX834949 2448 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 2073 from Patent EPI347046.
ACCESSION AX834949
VERSION AX834949.1 GI:39921084
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Yamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Maeho, Y.
Full-length cDNA sequences
Patent: EP 1347046-A 2073 24-SEP-2003;
Research Association for Biotechnology (JP)
Location/Qualifiers
1. 2448
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Query Match 35.2%; Score 944; DB 6; Length 2448;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 994; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 708 TGGGTCCTGGGGTGGTGGCATCTCCCTAAAGGAGTGGCTCTCTGGCAAGTGATGA 767
Db 357 TGGGTCCTGGGGTGGTGGCATCTCCCTAAAGGAGTGGCTCTCTGGCAAGTGATGA 416
QY 768 AGTCTCCAGCTTTTCCCTCAGCTCTCCAGACAGAAAGTGGCTCTCTGGTGGTGA 827
Db 417 AGTCTCCAGCTTTTCCCTCAGCTCTCCAGACAGAAAGTGGCTCTCTGGTGGTGA 476
QY 828 ACCCAGAGTACAGAGGATTTGGAGCCCGTGAAGAAAGTGGCTCTCTGGTGGTGA 887
Db 477 ACCCAGAGTACAGAGGATTTGGAGCCCGTGAAGAAAGTGGAGAGATGGGACCT 536
QY 888 TGACCTGAACGGGAGTGTGGTGGCAACACCCGTAGAAATGCCAAACCGTCCACGA 947
Db 537 TGACCTGAACGGGAGTGTGGTGGCAACACCCGTAGAAATGCCAAACCGTCCACGA 596
QY 948 GGAGCTCAGACAGCGCTGGAGAGCCAGACAGATGGAGAGAGCGTGCATGCATCAT 1007
Db 597 GGAGCTCAGACAGCGCTGGAGAGCCAGACAGATGGAGAGAGCGTGCATGCATCAT 656
QY 1008 CTGCGAGAGCTGCTGCAACGACTGCGTGGTGGTGGCAACCGTGCATGCATGCATGC 1067
Db 657 CTGCGAGAGCTGCTGCAACGACTGCGTGGTGGTGGCAACCGTGCATGCATGCATGC 716
QY 1068 GGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGCTCTACCTGCGCGCTGCCGT 1127
Db 717 GGCTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGCTCTACCTGCGCGCTGCCGT 776
QY 1128 GGAGCGGATCTGTAAACACCATCTTCAACACCTGCTGGAGCATACCTCATCCAGCA 1187
Db 777 GGAGCGGATCTGTAAACACCATCTTCAACACCTGCTGGAGCATACCTCATCCAGCA 836
QY 1188 TCCAGACAGAGTGGCAGTGAAGAGATGTGCAAGATATGGATGCCAGGATAAAATCAC 1247
Db 837 TCCAGACAGAGTGGCAGTGAAGAGATGTGCAAGATATGGATGCCAGGATAAAATCAC 896
QY 1248 TCAAGACATGCTGAGCGCCAAAGTCAAGCGGTCTTTTCTGATGAAGAGGATTCAGA 1307
Db 897 TCAAGACATGCTGAGCGCCAAAGTCAAGCGGTCTTTTCTGATGAAGAGGATTCAGA 956
QY 1308 GAGCCTGCTGAGCTGTGAGACGTTGACAGTGTGATGAGTGTGATGAGTGTGATGAGT 1367
Db 957 GAGCCTGCTGAGCTGTGAGACGTTGACAGTGTGATGAGTGTGATGAGTGTGATGAGT 1016
QY 1368 CGTGTGCGGAGTGTGCTGTGATGATGAGTGAAGAGCGCGCGGAGCTTCCCTGCCAGC 1427
Db 1017 CGTGTGCGGAGTGTGCTGTGATGATGAGTGAAGAGCGCGCGGAGCTTCCCTGCCAGC 1076
QY 1428 ACCGAGGGGAGCCAGGAGCCCAAGCGCTTGGGGATGTCACCTCCAGCTCCGTCAG 1487
Db 1077 ACCGAGGGGAGCCAGGAGCCCAAGCGCTTGGGGATGTCACCTCCAGCTCCGTCAG 1136

Qy	1488	CCTGACGACAGAGTCCAGGATTACGTGTGCCCTCTGCAAGGAAGCCACGCGCTGTGCAC	1547
Db	1137	CCTGACGACAGAGTCCAGGATTACGTGTGCCCTCTGCAAGGAAGCCACGCGCTGTGCAC	1196
Qy	1548	CTGTCGCTTCCAGAGCCCATGCCCCGACCGGAGCGCGAGCAGGACCCGCGCTGTGCG	1607
Db	1197	CTGTCGCTTCCAGAGCCCATGCCCCGACCGGAGTGGAGCGCAGCAGGACCGCGCTGTGCG	1256
Qy	1608	CCCTCAGCAGTGTGCGGTCTGCGCTCGACGCGCTTTCTGCCACCTGTACTGGGGCTGCACCCG	1667
Db	1257	CCCTCAGCAGTGTGCGGTCTGCGCTCGAGCCTTTCTGCCACCTGTACTGGGGCTGCACCCG	1316
Qy	1668	GACCGGCTGCTACGGCTGCTGTGCCCGCGTTTTGTG	1702
Db	1317	GACCGGCTGCTACGGCTGCTGTGCCCGCGTTTTGTG	1351

RESULT 12	AK097671.1	2448 bp	linear	PRI 15-JUL-2002
LOCUS	AK097671			
DEFINITION	Homo sapiens cDNA FLJ40352 fis, clone TEST12033505, highly similar to Homo sapiens cell cycle checkpoint protein CHFR mRNA.			
ACCESSION	AK097671			
VERSION	AK097671.1	GI:21757518		
KEYWORDS	oligo capping, fis (full insert sequence).			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			

	Query Match	35.2%	Score 944	DB 9	Length 2448
	Best Local Similarity	99.9%	Pred. NO. 0		
	Matches 994	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	708	TGGGTCCTGGGGGGTGTGGCATCTCCCTAAAGGAAGTGGTCCCTCTCTGCGAAGTATCA	767		
Db	357	TGGGTCCTGGGGGGTGTGGCATCTCCCTAAAGGAAGTGGTCCCTCTCTGCGAAGTATCA	416		

Qy	768	AGTCTCCAGCTTTGCCCTCAGCTCTCCACAGACAAAGACTGGTCCCTTTTGGTCGTTGGA	827
Db	417	AGTCTCCAGCTTTGCCCTCAGCTCTCCACAGACAAAGACTGGTCCCTTTTGGTCGTTGGA	476
Qy	828	ACCCACGAGTACAGAGGATTTGGAGCCCGTGAAGAAAGAAAATGAGAGGAGATGGGACCT	887
Db	477	ACCCACGAGTACAGAGGATTTGGAGCCCGTGAAGAAAGAAAATGAGAGGAGATGGGACCT	536
Qy	888	TGACCTGAAACGGGCAGTTGTGTGCTCCACAACCGCGTAGAAATGCCCAACCGTCCACGA	947
Db	537	TGACCTGAAACGGGCAGTTGTGTGCTCCACAACCGCGTAGAAATGCCCAACCGTCCACGA	596
Qy	948	GGACGTTCAGAGCAGCGGCTGGGAAGCCAGACAAGATGGAGGAGAGCGCTGACATGCATCAT	1007
Db	597	GGACGTTCAGAGCAGCGGCTGGGAAGCCAGACAAGATGGAGGAGAGCGCTGACATGCATCAT	656
Qy	1008	CTGCCAGGACCTGCTGTCACGACTGCGTGAAGTTGACGCCCTGCAATGACACACGTTCTGCGC	1067
Db	657	CTGCCAGGACCTGCTGTCACGACTGCGTGAAGTTGACGCCCTGCAATGACACACGTTCTGCGC	716
Qy	1068	GGCTTGCTACTCGGCGTGCATGAGAGCGCTGCTCCCTGTGTCTACTCTGCGGTGTCCTG	1127
Db	717	GGCTTGCTACTCGGCGTGCATGAGAGCGCTGCTCCCTGTGTCTACTCTGCGGTGTCCTG	776
Qy	1128	GGAGCGGATCTGTAAATAACCAATCTCTCAACAACTCGTGGAGGACTACTCTCATTCAGCA	1187
Db	777	GGAGCGGATCTGTAAATAACCAATCTCTCAACAACTCGTGGAGGACTACTCTCATTCAGCA	836
Qy	1188	TCAGACAGAGTTCGAGCTGAAGAGATGTGCAAAAGTATGGATGCGAGGAATAAAATCAC	1247
Db	837	TCAGACAGAGTTCGAGCTGAAGAGATGTGCAAAAGTATGGATGCGAGGAATAAAATCAC	896
Qy	1248	TCAAGACATGCTGCAGCCAAAGTACAGCGGCTTTTCTGATGAAGAGGAGTTTCAGA	1307
Db	897	TCAAGACATGCTGCAGCCAAAGTACAGCGGCTTTTCTGATGAAGAGGAGTTTCAGA	956
Qy	1308	GGACCTGCTGAGCTGTTCAGAGTTTGAAGTGAAGTCTTCAGACATATGACAGCCATACGT	1367
Db	957	GGACCTGCTGAGCTGTTCAGAGCTTTGACAGTGAAGTCTTCAGACATATGACAGCCATACGT	1016
Qy	1368	CGTGTCCGCGAGTGTCTGAGTACAGAGGACGAGCGCGCAGCTCCCACTTGCCACAG	1427
Db	1017	CGTGTCCGCGAGTGTCTGAGTACAGAGGACGAGCGCGCAGCTCCCACTTGCCACAG	1076
Qy	1428	ACCCAGGCGAGCCAGGAGCCCCAGGCCCTGGGGGATGACCCCTCCAGTCCGTTCAG	1487
Db	1077	ACCCAGGCGAGCCAGGAGCCCCAGGCCCTGGGGGATGACCCCTCCAGTCCGTTCAG	1136
Qy	1488	CCCTGACGACAGAGTTCAGGATTAAGTGTGCCCCCTGTGCAAGGAAGCAGCCCTGTGAC	1547
Db	1137	CCCTGACGACAGAGTTCAGGATTAAGTGTGCCCCCTGTGCAAGGAAGCAGCCCTGTGAC	1196
Qy	1548	GTGCTGCTTCCAGCCCATGCCGAGCCGAGAGCGAGCGCAGCAGGACCCCGTGTGCG	1607
Db	1197	GTGCTGCTTCCAGCCCATGCCGAGCCGAGAGTGGAGCGCAGCAGGACCCCGTGTGCG	1256
Qy	1608	CCCTCAGCAGTGTGGGCTCTGCTGCAGCTTTTCTGCAACCTGTACTGGGGCTGCACCG	1667
Db	1257	CCCTCAGCAGTGTGGGCTCTGCTGCAGCTTTTCTGCAACCTGTACTGGGGCTGCACCG	1316
Qy	1668	GAACCGCTGCTACGGCTGCTGCCCCCGTTTGTG	1702
Db	1317	GACCGGCTGCTACGGCTGCTGCCCCCGTTTGTG	1351

[illegible]

Qy 2276 GGGGGCCGAGCCCGCCCTCCTGGTGAGCGCTGGGCGAGGGCTCGTGGTCATCAGCAGC 2335
Db |||||||
Qy 1173 GGGGGCCGAGCCCGCCCTCCTGGTGAGCGCTGGGCGAGGGCTCGTGGTCATCAGCAGC 1232
Db |||||||
Qy 2336 AGAGACGAAGCCCTTCTGTAAACATGCGCCGTCCTCCGCGAGAGGGGCAAGTTTGTCTTT 2395
Db |||||||
Qy 1233 AGAGACGAAGCCCTTCTGTAAACATGCGCCGTCCTCCGCGAGAGGGGCAAGTTTGTCTTT 1292
Db |||||||
Qy 2396 TGTACATTTTCCGAARACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAGTTTCAAGG 2455
Db |||||||
Qy 1293 TGTACATTTTCCGAARACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAGTTTCAAGG 1352
Db |||||||
Qy 2456 AGAAGGGCAAGTTTATCAAAAACATTTGTTTCAGGAGAGGGAGCATAGTTTACAGCCTA 2515
Db |||||||
Qy 1353 AGAAGGGCAAGTTTATCAAAAACATTTGTTTCAGGAGAGGGAGCATAGTTTACAGCCTA 1412
Db |||||||
Qy 2516 CAGGACGTACACAATATCCTGCTGCTGGGAAACCAACAGCATTTTATCTATTTTATTT 2575
Db |||||||
Qy 1413 CAGGACGTACACAATATCCTGCTGCTGGGAAACCAACAGCATTTTATCTATTTTATTT 1472
Db |||||||
Qy 2576 TAATAGTTTGGTGCTTATCTTCTAATAAGATTTAAATGTCACAACTGTAGCAAAATA 2635
Db |||||||
Qy 1473 TAATAGTTTGGTGCTTATCTTCTAATAAGATTTAAATGTCACAACTGTAGCAAAATA 1532
Db |||||||
Qy 2636 ATATAATTTATAATTTACAAATTGAC 2661
Db |||||||
Qy 1533 ATATAATTTATAATTTACAAATTGAC 1558
Db |||||||

Search completed: May 15, 2004, 05:09:34
Job time : 16432 secs

Query Match		35.2%;	Score 944;	DB 16;	Length 2448;			
Best Local Similarity		99.9%;	Pred. No. 0;					
Matches		994;	Conservative	0;	Mismatches	1;	Indels	0; Gaps 0;
Qy	708	TGGGTCTGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAAGTATGA	767					
Db	357	TGGGTCTGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAAGTATGA	416					
Qy	768	AGTCTCCAGCTTTGGCTCAGCTCTCCACAGACAGAAAGACTCGCTCTCTTTTCGTGTTGA	827					
Db	417	AGTCTCCAGCTTTGGCTCAGCTCTCCACAGACAGAAAGACTCGCTCTCTTTTCGTGTTGA	476					
Qy	828	ACCCAGGATCAGAGAGATTGGAGCCCGTGAAGAGAAATGAGAGAGATGGGACCT	887					
Db	477	ACCCAGGATCAGAGAGATTGGAGCCCGTGAAGAGAAATGAGAGAGATGGGACCT	536					
Qy	888	TGACTCAACGGGAGTGTGGTTCGACAAACCGCTAGAAATGCCAAACCGTCCACGA	947					
Db	537	TGACTCAACGGGAGTGTGGTTCGACAAACCGCTAGAAATGCCAAACCGTCCACGA	596					
Qy	948	GGAGCTCAGAGACGCGCTGGGAAGCAGACAAAGATGGAGAGACGCTGACATGCATCAT	1007					
Db	597	GGAGCTCAGAGACGCGCTGGGAAGCAGACAAAGATGGAGAGACGCTGACATGCATCAT	656					
Qy	1008	CTGCCAGACCTGTGCACGACTCGGTGAGTGTGGAGCCCTGCATGCACACGTTTCGCC	1067					
Db	657	CTGCCAGACCTGTGCACGACTCGGTGAGTGTGGAGCCCTGCATGCACACGTTTCGCC	716					
Qy	1068	GGCTTGCTACTCGGGCTGGAGTGGAGCGCTCGTCCCTGTGTCTTACCTGCGCTGTCCCGT	1127					
Db	717	GGCTTGCTACTCGGGCTGGAGTGGAGCGCTCGTCCCTGTGTCTTACCTGCGCTGTCCCGT	776					
Qy	1128	GGAGCGATCTGTMAAACACATCTCTCAACACCTCGTGGAGAGCATACCTCATCCAGCA	1187					
Db	777	GGAGCGATCTGTMAAACACATCTCTCAACACCTCGTGGAGAGCATACCTCATCCAGCA	836					
Qy	1188	TCCAGACAAGATGCGAGTGAAGAAGATGTGCAAGATGATGGATGCCAGGAATAAAATCAC	1247					
Db	837	TCCAGACAAGATGCGAGTGAAGAAGATGTGCAAGATGATGGATGCCAGGAATAAAATCAC	896					
Qy	1248	TCAAGACATGCTGCAGCCCAAGTCAAGCGCTCTTTTCTGATGAAGAGGAGTTCAGA	1307					
Db	897	TCAAGACATGCTGCAGCCCAAGTCAAGCGCTCTTTTCTGATGAAGAGGAGTTCAGA	956					
Qy	1308	GGACTGCTGAGCTGTGACAGCTTGACAGTGTCTCAGACATTAGCCAGGCCATACGT	1367					
Db	957	GGACTGCTGAGCTGTGACAGCTTGACAGTGTCTCAGACATTAGCCAGGCCATACGT	1016					
Qy	1368	CGTGTCCGGCAGTGTCTGAGTACAGAGGCGGCGGCGCAGCTCCCTCCCTGCTCCAGC	1427					
Db	1017	CGTGTCCGGCAGTGTCTGAGTACAGAGGCGGCGGCGCAGCTCCCTCCCTGCTCCAGC	1076					
Qy	1428	ACCCGAGGCGGCGCAGGAGCCCAAGGCGCTGGGGGATGCACCTTCCACGTCGCTCAG	1487					
Db	1077	ACCCGAGGCGGCGCAGGAGCCCAAGGCGCTGGGGGATGCACCTTCCACGTCGCTCAG	1136					
Qy	1488	CCTGACACAGCAGTCCAGGATTAAGTGTGCCCTCTGCAAGAGGAGCCAGCCCTGTGCAC	1547					
Db	1137	CCTGACACAGCAGTCCAGGATTAAGTGTGCCCTCTGCAAGAGGAGCCAGCCCTGTGCAC	1196					
Qy	1548	CTGTGCTTCCAGCCATGCGGACCGGAGAGCGGAGCGGAGACCGCGGTGTCG	1607					
Db	1197	CTGTGCTTCCAGCCATGCGGACCGGAGAGTGAGCGGAGCGGAGACCGCGGTGTCG	1256					
Qy	1608	CCCTCAGCAGTGTGCGGTCTGCTGACGCTTCTTCTGACCTGTACTTGGGGCTGCACCCG	1667					
Db	1257	CCCTCAGCAGTGTGCGGTCTGCTGACGCTTCTTCTGACCTGTACTTGGGGCTGCACCCG	1316					
Qy	1668	GACCGGTGCTACGGCTGCTGCGCCCGCTTTTGTG	1702					
Db	1317	GACCGGTGCTACGGCTGCTGCGCCCGCTTTTGTG	1351					

Query Match		28.6%;	Score 765;	DB 9;	Length 1311;			
Best Local Similarity		100.0%;	Pred. No. 0;					
Matches		765;	Conservative	0;	Mismatches	0;	Indels	0; Gaps 0;
Qy	710	GGTCTGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAAGTATGAAG	769					
Db	383	GGTCTGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAAGTATGAAG	442					
Qy	770	TCTCCAGCTTTGCCCTCAGCTCTCCAGACAGAAAGACTCGCTCTTTTCGTTGGAAAC	829					
Db	443	TCTCCAGCTTTGCCCTCAGCTCTCCAGACAGAAAGACTCGCTCTTTTCGTTGGAAAC	502					
Qy	830	CCAGGATCAGAGAGATTGGAGCCCGTGAAGAGAAATGAGAGAGATGGGGACCTTG	889					
Db	503	CCAGGATCAGAGAGATTGGAGCCCGTGAAGAGAAATGAGAGAGATGGGGACCTTG	562					
Qy	890	ACCTCAACGGGAGTGTGGTTCGCACAAACCGCTAGAAATGCCAAACCGTCCACGAGG	949					
Db	563	ACCTCAACGGGAGTGTGGTTCGCACAAACCGCTAGAAATGCCAAACCGTCCACGAGG	622					
Qy	950	ACGTGAGAGACGCGCTGGGAAGCAGACAAAGATGGAGAGACGCTGACATGCATCATCT	1009					
Db	623	ACGTGAGAGACGCGCTGGGAAGCAGACAAAGATGGAGAGACGCTGACATGCATCATCT	682					
Qy	1010	GCCAGACCTGCTGCACGACTCGCTGAGTGTGGAGCCCTGCTGACACACGTTCTGCGCGG	1069					
Db	683	GCCAGACCTGCTGCACGACTCGCTGAGTGTGGAGCCCTGCTGACACACGTTCTGCGCGG	742					
Qy	1070	CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCTTACCTGCGCTGTCCCGTGG	1129					
Db	743	CTTGCTACTCGGGCTGGATGGAGCGCTCGTCCCTGTGTCTTACCTGCGCTGTCCCGTGG	802					
Qy	1130	ACCGGATCTGTAAACACCATCTTCAACACCTCGTGGAGAGCATACCTCATCCAGCATC	1189					
Db	803	ACCGGATCTGTAAACACCATCTTCAACACCTCGTGGAGAGCATACCTCATCCAGCATC	862					
Qy	1190	CAGACAGAGTCCGAGTGAAGAAGATGTGCAAAAGTATGGATGCCAGGAATAAAATCACTC	1249					
Db	863	CAGACAGAGTCCGAGTGAAGAAGATGTGCAAAAGTATGGATGCCAGGAATAAAATCACTC	922					
Qy	1250	AAGACATGCTGAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAGGGAGTTCAGAGG	1309					
Db	923	AAGACATGCTGAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAGGGAGTTCAGAGG	982					
Qy	1310	ACCTGCTGAGCTGTGACAGCTTGAAGTGTGAGTCTCAGACATTAGCCAGGACCATACGTCG	1369					
Db	983	ACCTGCTGAGCTGTGACAGCTTGAAGTGTGAGTCTCAGACATTAGCCAGGACCATACGTCG	1042					
Qy	1370	TGTGCGGCGAGTGTCTTGAGTACAGAGCGCGCGCGCTCCCACTGCGCCAGGAC	1429					

Db 1043 TGTGCGGCGAGTCTCTGAGTACAGAGGCGGCGGCGAGCCTCCCACTGCGCCAGCAC 1102
Qy 1430 CCGAGGGCGGCGAGGAGCCCAAGAGCCCTGCGGGGATGACACCT 1474
Db 1103 CCGAGGGCGGCGAGGAGCCCAAGAGCCCTGCGGGGATGACACCT 1147

RESULT 4
US-10-094-749-160
; Sequence 160, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUJIKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KACRU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKI, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTOKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 160
; LENGTH: 2186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-160

Query Match 23.6%; Score 633; DB 16; Length 2186;
Best Local Similarity 99.7%; Pred. No. 2e-308;
Matches 863; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 1796 TCACATGGAACATGTTGACCGAGGCGCTCGTGGCTCCAGCGGGAGTGTCTGCG 1855
Db 695 TGACATGGAACATGTTGACCGAGGCGCTCGTGGCTCCAGCGGGAGTGTCTGCG 754

Qy 1856 TGTCTGATTACAGAGTCAAGGAGACACCGTTCTGTGTTACTGCTGTGGCTCGCGAGCT 1915
Db 755 TGTCTGATTACAGAGTCAAGGAGACACCGTTCTGTGTTACTGCTGTGGCTCGCGAGCT 814

Qy 1916 TCCGTGAGCTACCTATCAGTATCGGAGACATTCCTGCTCCGAGTTCGAGTGGCGG 1975
Db 815 TCCGTGAGCTACCTATCAGTATCGGAGACATTCCTGCTCCGAGTTCGAGTGGCGG 874

Qy 1976 TAACATCCGCTCCTGACTGCTACTGCGGCGGTAACTGCCGCACTCAGGTGAAAGCTCACC 2035
Db 875 TAACATCCGCTCCTGACTGCTACTGCGGCGGTAACTGCCGCACTCAGGTGAAAGCTCACC 934

Qy 2036 ACGGCATGAATTCATCATATCTGTGAACAGACAGGTTCAAAACTAAGCATCCAGAG 2095
Db 935 ACGGCATGAATTCATCATATCTGTGAACAGACAGGTTCAAAACTAAGCATCCAGAG 994

Qy 2096 GCCCTGACAGCTTTTACAGCTGGAGGTGAAGAGAGCGTGTGTTTAAATACAGAGACAA 2155
Db 995 GCCCTGACAGCTTTTACAGCTGGAGGTGAAGAGAGCGTGTGTTTAAATACAGAGACAA 1054

Qy 2156 GCAGTCAAGGTGTTTTCACAGCCCTGAGGAAGGAGCGAGGCTCTCCGACAGGTGC 2215
Db 1055 ACAGTCAAGGTGTTTTCACAGCCCTGAGGAAGGAGCGAGGCTCTCCGACAGGTGC 1114
Qy 2216 TCTGGGTGACTCTTCTGTGGAGCTTTTACCCCTCTCTGAGTACACCTCCCCAGAGCCCC 2275
Db 1115 TCTGGGTGACTCTTCTGTGGAGC--TTTACCCTCTCTGAGTACACCTCCCCAGAGCCCC 1172
Qy 2276 GGGGGCGGAGCGCCCTCTGAGTGGAGGCTGGGAGGAGGCTGCTGGTGGCATCAGCAGC 2335
Db 1173 GGGGGCGGAGCGCCCTCTGAGTGGAGGCTGGGAGGAGGCTGCTGGTGGCATCAGCAGC 1232
Qy 2336 AGAGACGAAGCCTTTCTGTAACTGCGGCGCTCCCGGAGAGGAGGCTTTTGTCTCTTT 2395
Db 1233 AGAGACGAAGCCTTTCTGTAACTGCGGCGCTCCCGGAGAGGAGGCTTTTGTCTCTTT 1292
Qy 2396 TGTACATTTCCGAAACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAAGTTTCAAGGG 2455
Db 1293 TGTACATTTCCGAAACTACAGTTAAAGCAGAGTCTGTTTTCAGGAAAAGTTTCAAGGG 1352
Qy 2456 AGAAGGCAAGTTTATCAAAAACATTTGTTTCAGGAGAAAGGAGCATTAAGTTACAGCCTA 2515
Db 1353 AGAAGGCAAGTTTATCAAAAACATTTGTTTTCAGGAGAAAGGAGCATTAAGTTACAGCCTA 1412
Qy 2516 CAGGACGTACACATATCTCTGCTGGGAAAACACACGATTTTATTTTATTTT 2575
Db 1413 CAGGACGTACACATATCTCTGCTGGGAAAACACACGATTTTATTTTATTTT 1472
Qy 2576 TAATAGTTTGGTCTTATCTTCTAATAAGATTTAAATGTCAAACTGTAGCAGCAATA 2635
Db 1473 TAATAGTTTGGTCTTATCTTCTAATAAGATTTAAATGTCAAACTGTAGCAGCAATA 1532
Qy 2636 ATATAATTTTATTAATTACAAATTTGAC 2661
Db 1533 ATATAATTTTATTAATTACAAATTTGAC 1558

RESULT 5
US-09-764-864-483
; Sequence 483, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 483
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (623)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (640)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (670)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (687)

OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (690)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-483

Query Match 17.0%; Score 455; DB 9; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.3e-218; Mismatches 0; Indels 0; Gaps 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1044 GCCTGATGACACACGTTCTGCGGGTTCCTACTGCGGTGATGAGCGCTCGTCCCT 1103
DB 14 GCCCTGATGACACACGTTCTGCGGGTTCCTACTGCGGTGATGAGCGCTCGTCCCT 73

QY 1104 GTGTCCTACCTCCGCTGTCCTGCGGAGCGATCTGTAAAAACCAATCTCAACAACT 1163
DB 74 GTGTCCTACCTCCGCTGTCCTGCGGAGCGATCTGTAAAAACCAATCTCAACAACT 133

QY 1164 CGTGGAGCATACCTCATCCAGCATCCAGACAGAGTCGAGTGAAGAGATGCAAG 1223
DB 134 CGTGGAGCATACCTCATCCAGCATCCAGACAGAGTCGAGTGAAGAGATGCAAG 193

QY 1224 TATGATGCCAGGATTAATCACTCAAGCATCTGACGCCCAAGTCAAGCGGTCTTT 1283
DB 194 TATGATGCCAGGATTAATCACTCAAGCATCTGACGCCCAAGTCAAGCGGTCTTT 253

QY 1284 TTCTGATGAAGAAGGAGTTTCAGAGGACCTGCTGGAGCTGTCAAGCTTCAGAGTGC 1343
DB 254 TTCTGATGAAGAAGGAGTTTCAGAGGACCTGCTGGAGCTGTCAAGCTTCAGAGTGC 313

QY 1344 CTCAGACATTAGCCAGCATACGTCGTGTCGCGGAGTCTCTGAGTACAGAGGCGGC 1403
DB 314 CTCAGACATTAGCCAGCATACGTCGTGTCGCGGAGTCTCTGAGTACAGAGGCGGC 373

QY 1404 GGCAGCCTCCCTCCAGTCCAGCACCGGAGGCGAGCCAGGAGCCCAAGGCGCTGGG 1463
DB 374 GGCAGCCTCCCTCCAGTCCAGCACCGGAGGCGAGCCAGGAGCCCAAGGCGCTGGG 433

QY 1464 GGATGACCTCCAGTCCGCTGAGCCTGACGACAG 1498
DB 434 GGATGACCTCCAGTCCGCTGAGCCTGACGACAG 468

RESULT 6
US-09-764-864-322
; Sequence 322, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 322
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (272)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (300)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-322

Query Match 10.0%; Score 267; DB 9; Length 357;
Best Local Similarity 100.0%; Pred. No. 9.9e-124; Mismatches 0; Indels 0; Gaps 0;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7
US-09-764-864-323
; Sequence 323, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 323
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (289)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (313)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-323

Query Match 8.9%; Score 238; DB 9; Length 354;
Best Local Similarity 100.0%; Pred. No. 4.3e-109; Mismatches 0; Indels 0; Gaps 0;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 GCAATAAACTGCTCTGAGATCACTGTAGAAATGTAGTGAATAAATCAGGTCAGG 304
DB 51 GCATAAACTGCTCTGAGATCACTGTAGAAATGTAGTGAATAAATCAGGTCAGG 110

QY 305 TGACACTGGAAGATACCAAGCAGTGAACAGTGAATTAACAGCTGAAGTTGTTAGA 364
DB 111 TGACACTGGAAGATACCAAGCAGTGAACAGTGAATTAACAGCTGAAGTTGTTAGA 170

QY 365 AGCAGACATGCCCTTTACAGACTGGGATGTCTACTTGGGTACAGAAAGATGAAC 424
DB 171 AGCAGACATGCCCTTTACAGACTGGGATGTCTACTTGGGTACAGAAAGATGAAC 230

QY 425 CGGAACACACGTCGATACCTCTATGAATCTTTAAGTGAAGCAAGCATGACACA 482
DB 231 CGGAACACACGTCGATACCTCTATGAATCTTTAAGTGAAGCAAGCATGACACA 288

RESULT 8
US-09-764-864-742
; Sequence 742, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT23
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 742
LENGTH: 354
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (289)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (313)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-742

Query Match 8.9%; Score 238; DB 9; Length 354;
Best Local Similarity 100.0%; Pred. No. 4.3e-109;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 GCAATAAAGCTGCTCTGGAGATCACTGTAGATTGTTAGTGAATGATGAAAAATCAGTCAAG 304
DB 51 GCAATAAAGCTGCTCTGGAGATCACTGTAGATTGTTAGTGAATGATGAAAAATCAGTCAAG 110
QY 305 TGACACTGGAAGATACACAGCAGTGGAAACAGTGAATTAACAGCTGAAGTGTGTAAGA 364
DB 111 TGACACTGGAAGATACACAGCAGTGGAAACAGTGAATTAACAGCTGAAGTGTGTAAGA 170
QY 365 AGCAGATGCGCTTTACAGATCGGGATGTCATCTACTTGTGTACAGGAGATGAAC 424
DB 171 AGCAGATGCGCTTTACAGATCGGGATGTCATCTACTTGTGTACAGGAGATGAAC 230
QY 425 CGGAACACACAGTGGATACCTCTATGAATCTTTAAGTGAAAAAGCAAGGCATGACACA 482
DB 231 CGGAACACACAGTGGATACCTCTATGAATCTTTAAGTGAAAAAGCAAGGCATGACACA 288

RESULT 9
US-09-918-995-13662
Sequence 13662, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13662
LENGTH: 449
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(449)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-13662

Query Match 8.1%; Score 217; DB 10; Length 449;
Best Local Similarity 100.0%; Pred. No. 1.7e-98;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2445 AGTTTCAAGGAGAGGCGAAGTTTATCAAAACATGTTTTCAGGAGAGGCGATAG 2504

DB 92 AGTTTCAAGGAGAGGCGAAGTTTATCAAAACATGTTTTCAGGAGAGGCGATAG 151
QY 2505 TTTACAGCCTTACAGGAGCTACACAAATATCTGCTGCTGGAAAAACACAGCATTTTATCT 2564
DB 152 TTTACAGCCTTACAGGAGCTACACAAATATCTGCTGCTGGAAAAACACAGCATTTTATCT 211
QY 2565 ATTTTATTTTAAATAGGTTTGGTCTTATCTTCTAATAGATTTAAATGTCAAAACTG 2624
DB 212 ATTTTATTTTAAATAGGTTTGGTCTTATCTTCTAATAGATTTAAATGTCAAAACTG 271
QY 2625 TAGCACAATATATATATTTATTAATTTTACAAATGAC 2661
DB 272 TAGCACAATATATATATTTATTAATTTTACAAATGAC 308

RESULT 10
US-09-918-995-2180
Sequence 2180, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2180
LENGTH: 476
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(476)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2180

Query Match 2.9%; Score 77; DB 10; Length 476;
Best Local Similarity 100.0%; Pred. No. 7.5e-28;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2046 ATTCATCATATCTGTGAACAGCAAGGTTCAAAACTAAGCATCCAGAGCCCTGAGCA 2105
DB 52 ATTCATCATATCTGTGAACAGCAAGGTTCAAAACTAAGCATCCAGAGCCCTGAGCA 111
QY 2106 GCTTTCAGCACTGGAGG 2122
DB 112 GCTTTCAGCACTGGAGG 128

RESULT 11
US-09-908-975-6220
Sequence 6220, Application US/09908975
Publication No. US20030165843A1
GENERAL INFORMATION:
APPLICANT: SHOSHAN, Avi
APPLICANT: WASSERMAN, Alon
APPLICANT: MINTZ, Eli
APPLICANT: MINTZ, Liat
APPLICANT: FAIGLER, Simchon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6220
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-6220

Query Match      2.2%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.8e-19;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1843 GGAGTGTTCGTCTGTTACAGAGTCACGGGAGACACCGTCTGTGTTACTGCTGT 1902
Db 1 GGAGTGTTCGTCTGTTACAGAGTCACGGGAGACACCGTCTGTGTTACTGCTGT 60

RESULT 12
US-10-304-928-15
; Sequence 15, Application US/10304928
; Publication No. US20030101484A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Yalpani, Naaser
; TITLE OF INVENTION: Maize Chitinases and Their Use in
; TITLE OF INVENTION: Enhancing Disease Resistance in Crop Plants
; FILE REFERENCE: 1100D
; CURRENT APPLICATION NUMBER: US/10/304,928
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/522,714
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/125,915
; PRIOR FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 1085
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(966)
US-10-304-928-15

Query Match      0.9%; Score 25; DB 15; Length 1085;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679
Db 1061 AATTGACAAAAA 1085

RESULT 13
US-09-957-664-1
; Sequence 1, Application US/09957664
; Patent No. US20020123097A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; TITLE OF INVENTION: A NOVEL HUMAN TRANSPORTER AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-192
; CURRENT APPLICATION NUMBER: US/09/957,664
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,790
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1963
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (67)...(1491)
US-09-957-664-1

Query Match      0.9%; Score 25; DB 9; Length 1963;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679
Db 1916 AATTGACAAAAA 1940

RESULT 14
US-10-154-419-15
; Sequence 15, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62052, 8099, 46455, 54144, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FHS8295FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 15
; LENGTH: 1963
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(1491)
US-10-154-419-15

Query Match      0.9%; Score 25; DB 15; Length 1963;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679
Db 1916 AATTGACAAAAA 1940

RESULT 15
US-09-938-842A-3337/c
; Sequence 3337, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3337
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```

US-09-938-842A-3337

Query Match 0.9%; Score 25; DB 9; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679
DB 1268 AATTGACAAAAA 1244

Search completed: May 15, 2004, 12:03:05
Job time : 1595 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2004, 23:36:46 ; Search time 10039 Seconds
(without alignments)
7968.999 Million cell updates/sec

Title: US-10-048-046-1

Perfect score: 2679

Sequence: 1 asgaattgcgcagggccg.....acaaaaaaaaaaaaaaaaa 2679

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estnu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_psq:*
 - 27: em_gss_vrl:*
 - 28: gp_gss1:*
 - 29: gp_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825	30.8	1201	9	AL539798
2	722	27.0	921	13	BUI176559
3	703	26.2	992	13	BQ073354
4	698	26.1	877	14	CD243458

5	647	24.2	897	10	BF971929
6	620	23.1	840	12	BG395714
7	600	22.4	680	9	AUI134937
8	583	21.8	870	13	BQ215731
9	565	21.1	915	12	BM459647
10	563	21.0	1143	12	BM542545
11	560	20.9	666	10	BE407732
12	539	20.1	1201	9	AL539797
13	519	19.4	934	10	BF971838
14	513	19.1	729	14	CA916737
15	488	18.2	690	10	BE280165
16	478	17.8	910	13	BUI148471
17	478	17.8	938	13	BQ887242
18	477	17.8	1092	12	BM541730
19	469	17.5	909	13	EX409791
20	468	17.5	824	9	AUI132825
21	456	17.0	712	10	BE257482
22	456	17.0	1095	12	BG335362
23	452	16.9	1022	12	BM558904
24	450	16.8	953	13	BQ936072
25	447	16.7	938	13	BQ924736
26	436	16.3	535	13	EX486289
27	436	16.3	557	13	EX486564
28	435	16.2	811	14	CB989147
29	433	16.2	806	14	CB961597
30	426	15.9	518	9	AUI54212
31	421	15.7	687	9	AUI138310
32	420	15.7	876	13	BQ876252
33	418	15.6	650	10	AW957709
34	412	15.4	635	12	BG481936
35	409	15.3	490	10	BF109790
36	404	15.1	881	13	BUI153737
37	394	14.7	944	13	EX409790
38	386	14.4	475	9	AI200393
39	382	14.3	824	12	BG831289
40	381	14.2	797	12	BI460002
41	364	13.6	1070	12	BM802324
42	354	13.2	636	10	BE407797
43	335	12.5	697	12	BG473597
44	326	12.2	527	10	BF439206
45	320	11.9	722	9	AUI135780

ALIGNMENTS

RESULT 1
AL539798
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL539798 1201 bp mRNA linear EST 31-MAY-2003
AL539798 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSODF031YA23 5-PRIME, mRNA sequence.
AL539798
EST.
AL539798.2 GI:31264361
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
Li W.B., Gruber C., Jesse J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12869349.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6792.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODF031YA23&cluster=6792.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF031A12QF1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF031A23"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

FEATURES
source

ORIGIN

Query Match 30.8%; Score 825; DB 9; Length 1201;
Best Local Similarity 100.0%; Pred. No. 1.4e-155;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GTTCGGCGGGCGGGGATGTAATCCGATGGAGCGCGCGGAGGCAAGCAGTCG 120
DB 67 GTTCGGCGGGCGGGGATGTAATCCGATGGAGCGCGCGGAGGCAAGCAGTCG 126

QY 121 CGCCGCGCGACCTCTGGGACGGCTCTCGCTCTGGGCGGGGAGGCGAGCGCGAC 180
DB 127 CGCCGCGCGACCTCTGGGACGGCTCTCGCTCTGGGCGGGGAGGCGAGCGCGAC 186

QY 181 GTCTCTCTGAGAACGGGAGTGACATCGGGCGGAGACAGGTTGGACCTTCTTC 240
DB 187 GTCTCTCTGAGAACGGGAGTGACATCGGGCGGAGACAGGTTGGACCTTCTTC 246

QY 241 CCAGCAATAACTGTCTCTGAGATCACTAGATTTAGTGGATGAAATTCAGGT 300
DB 247 CCAGCAATAACTGTCTCTGAGATCACTAGATTTAGTGGATGAAATTCAGGT 306

QY 301 CAGGTGACACTGGAAGATACCAAGCCAGTGGAACAGTGATTAACAAGCTGAAGTTGTT 360
DB 307 CAGGTGACACTGGAAGATACCAAGCCAGTGGAACAGTGATTAACAAGCTGAAGTTGTT 366

QY 361 AAGAGCAGACATGCCCTTACAGACTGGGATGTCATCTCTGCTGTACAGGAAGAT 420
DB 367 AAGAGCAGACATGCCCTTACAGACTGGGATGTCATCTCTGCTGTACAGGAAGAT 426

QY 421 GAACGGGAACACAGCTGGCATACCTCTATGATCTTTAAGTGAAGCAAGGCGATGCA 480
DB 427 GAACGGGAACACAGCTGGCATACCTCTATGATCTTTAAGTGAAGCAAGGCGATGCA 486

QY 481 CAAGATCTTTGAAGCTAACAGGAATGTTTCCATGGACCAAGATACCTCAGGT 540
DB 487 CAAGATCTTTGAAGCTAACAGGAATGTTTCCATGGACCAAGATACCTCAGGT 546

QY 541 GCAGGTGACGGGCGAGGGCGGATCCCGGGTCCCTCGCTCGTCCCGCCACTCAGGTG 600
DB 547 GCAGGTGACGGGCGAGGGCGGATCCCGGGTCCCTCGCTCGTCCCGCCACTCAGGTG 606

QY 601 TGTCTTGAGGAACCAAGCCATCAATCGAGTCAGACCTTCCCGCCAGCCCTCGGCC 660
DB 607 TGTCTTGAGGAACCAAGCCATCAATCGAGTCAGACCTTCCCGCCAGCCCTCGGCC 666

QY 661 TCTTCCACGGAGCTTCTCTGCGAGGGCGAGAGCGTTCTCCAGTTGTGGGTCTGGGGGT 720
DB 667 TCTTCCACGGAGCTTCTCTGCGAGGGCGAGAGCGTTCTCCAGTTGTGGGTCTGGGGGT 726

QY 721 GGTGGCATCTCCCTAAAGAGTGTCTCTGTCGAGGCGAGAGCGTTCTCCAGTTGTGGGTCTGGGGGT 780
DB 727 GGTGGCATCTCCCTAAAGAGTGTCTCTGTCGAGGCGAGAGCGTTCTCCAGTTGTGGGTCTGGGGGT 786

QY 781 GCTCAGCTCTCCAGAGAGAGCTGGTCTCTTTCTGCTGTAAGCCCGCAGGATCAG 840
DB 787 GCTCAGCTCTCCAGAGAGAGCTGGTCTCTTTCTGCTGTAAGCCCGCAGGATCAG 846

QY 841 GAGGATTTGAGCCCGTGAAGAGAAATGAGAGAGATGGGGAC 885
DB 847 GAGGATTTGAGCCCGTGAAGAGAAATGAGAGAGATGGGGAC 891

RESULT 2
BUI76559
LOCUS
DEFINITION
BUI76559 921 bp mRNA linear EST 04-SRP-2002
AGENCOURT 7940317 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143609
5', mRNA sequence.
BUI76559
BUI76559.1 GI:22690543
VERSION
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 921)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL3467 row: e column: 18
High quality sequence stop: 697.
Location/Qualifiers
1..921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6143609"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

FEATURES
source

ORIGIN

Query Match 27.0%; Score 722; DB 13; Length 921;
Best Local Similarity 99.9%; Pred. No. 5.5e-135;
Matches 772; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1889 TGTGTTACTGCTGCGCTGCGCAGCTTCGTCAGCTGACCTATCAGTATCGGCAACA 1948
DB 1 TGTGTTACTGCTGCGCTGCGCAGCTTCGTCAGCTGACCTATCAGTATCGGCAACA 60

QY 1949 TTCTGTTCCGAGTTGCCAGTGGCGGTAACTCCCGTCTGACTGCTACTGGGCGGTA 2008
DB 61 TTCTGTTCCGAGTTGCCAGTGGCGGTAACTCCCGTCTGACTGCTACTGGGCGGTA 120

QY 2009 ACTGCCGACTCAGGTGAAAGCTCACAGCCATGAATTCATATCTGTGAACAGA 2068
DB 121 ACTGCCGACTCAGGTGAAAGCTCACAGCCATGAATTCATATCTGTGAACAGA 180

QY 2069 CAAGGTTCAAAACTAAGCATCCAGAGGCGCTTTCAGCAGCTTTCAGGAGGTGAGA 2128
DB 181 CAAGGTTCAAAACTAAGCATCCAGAGGCGCTTTCAGCAGCTTTCAGGAGGTGAGA 240

QY 2129 GAGCGTGTCTTTAAATACAGACAGCAGCTCAGGTGTTTTCACAGCCCTCCTGAGG 2188
DB 241 GAGCGTGTCTTTAAATACAGACAGCAGCTCAGGTGTTTTCACAGCCCTCCTGAGG 300

QY 2189 AAGGAGCGCAGGCTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGCTTTTACC 2248

```
Db 301 AAGGAGCGAGGCTCCGACAGGTCTGCGGTGAGCTCTCTGCGAGCTTTTACCC 360
Qy TCTGAGTGAGACCTCCCGAGAGCCCGGGGGCGCGAGCCCGCCCTCTGCTGAGCGCTG 2308
Db TCTGAGTGAGACCTCCCGAGAGCCCGGGGGCGCGAGCCCGCCCTCTGCTGAGCGCTG 420
Qy GCGAGGCTCTGCTGTCATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2368
Db GCGAGGCTCTGCTGTCATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy CCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2428
Db CCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy GTCTGTTTTTTCAGGAAAGATTTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2488
Db GTCTGTTTTTTCAGGAAAGATTTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2548
Db GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy CCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2608
Db CCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy TAAATGTCACAACTGTAGCACAATATATATATATATATATATATATATATATATAT 2661
Db TAAATGTCACAACTGTAGCACAATATATATATATATATATATATATATATATATAT 773
```

```
RESULT 3
BQ073354 992 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT_7052199 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:5806143
DEFINITION 5', mRNA sequence.
ACCESSION BQ073354
VERSION BQ073354.1 GI:19902400
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 992)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHC2046 row: h column: 16
High quality sequence stop: 670.
Location/Qualifiers
1. 992
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5806143"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
```

```
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
ORIGIN
Query Match 26.2%; Score 703; DB 13; Length 992;
Best Local Similarity 100.0%; Pred. No. 3.1e-131;
Matches 703; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 551 GCGCAGGGGGCGGATCCCGGGTCCCTCGTGTGCGCCCGCCACTCAGGTGTGCTTTGAGG 610
Db 1 GCGCAGGGGGCGGATCCCGGGTCCCTCGTGTGCGCCCGCCACTCAGGTGTGCTTTGAGG 60
Qy 611 AACACAGAGCCATCAACATCGAGCTCAGACCTCTTCCCCACAGCCCTCGGCCCTCTTCCACGG 670
Db 61 AACACAGAGCCATCAACATCGAGCTCAGACCTCTTCCCCACAGCCCTCGGCCCTCTTCCACGG 120
Qy 671 AGCCTTCTCTGAGGGGAGAGCGTTCTCTCAGTGTGGGTCTGGGTGTGGTGTGCACT 730
Db 121 AGCCTTCTCTGAGGGGAGAGCGTTCTCTCAGTGTGGGTCTGGGTGTGGTGTGCACT 180
Qy 731 CCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTGTCTCCAGCTTTGCCTCAGCTC 790
Db 181 CCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTGTCTCCAGCTTTGCCTCAGCTC 240
Qy 791 TCCACAGACAGAAAGACTGGTCTCTTTTCTGTGGAAACCCACAGGATCAGAGGATTTGG 850
Db 241 TCCACAGACAGAAAGACTGGTCTCTTTTCTGTGGAAACCCACAGGATCAGAGGATTTGG 300
Qy 851 AGCCCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 910
Db 301 AGCCCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 911 TCGCACAAACCGGTAGAAATGCCAAACCGTCCACAGAGAGAGTCCAGAGCAGCGGTGGGA 970
Db 361 TCGCACAAACCGGTAGAAATGCCAAACCGTCCACAGAGAGAGTCCAGAGCAGCGGTGGGA 420
Qy 971 AGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACT 1030
Db 421 AGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACT 480
Qy 1031 GGGTGTGAGTTCAGAGCCCTGATGACAGCTTCTGCGCGCTTGTCTACTCGGGCTGGATGG 1090
Db 481 GGGTGTGAGTTCAGAGCCCTGATGACAGCTTCTGCGCGCTTGTCTACTCGGGCTGGATGG 540
Qy 1091 AGCGTCTGCTCTGTGCTCTACCTGCCCTGTCCGTGTGGAGGAGATCTGTAAAAACCA 1150
Db 541 AGCGTCTGCTCTGTGCTCTACCTGCCCTGTCCGTGTGGAGGAGATCTGTAAAAACCA 600
Qy 1151 TCCTCAACAACTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1210
Db 601 TCCTCAACAACTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 1211 AAGATGTGCAAGTATGGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1253
Db 661 AAGATGTGCAAGTATGGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 703
```

```
RESULT 4
CD243458 877 bp mRNA linear EST 22-MAY-2003
LOCUS AGENCOURT_14121909 NIH_MGC_180 Homo sapiens cDNA clone
DEFINITION IMAGE:30383176 5', mRNA sequence.
ACCESSION CD243458
VERSION CD243458.1 GI:31003922
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 877)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```


JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
cDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM446 row: j column: 17
High quality sequence stop: 660.
Location/Qualifiers
1. 877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30383176"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH MGC 180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;
Site 2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

FEATURES

source

Query Match 26.1%; Score 698; DB 14; Length 877;
Best Local Similarity 99.9%; Pred. No. 3.4e-130;
Matches 748; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1527 AGAAGCCACGCGCTGTGACCTGTCTTCAGGCCATGCCGACCGAGAGCGGCG 1586
DB 4 AGAAGCCACGCGCTGTGACCTGTCTTCAGGCCATGCCGACCGAGAGCGGCG 63

QY 1587 CGAGCAGACCGCGCTGTGCGCCCTCAGAGTGTGGGTCTCCCTGAGCGCTTTCTGCCA 1646
DB 64 CGAGCAGACCGCGCTGTGCGCCCTCAGAGTGTGGGTCTCCCTGAGCGCTTTCTGCCA 123

QY 1647 CTTGTACTGGGCTGCACCGGACCGCTGTACGCTGCGCGCCGCTTTTGTGAGCT 1706
DB 124 CTTGTACTGGGCTGCACCGGACCGCTGTACGCTGCGCGCCGCTTTTGTGAGCT 183

QY 1707 CAACTGGGTGACAGTGTCTGAGCGGCTGTGAAACAAACACAGCTACGATCAGACAT 1766
DB 184 CAACTGGGTGACAGTGTCTGAGCGGCTGTGAAACAAACACAGCTACGATCAGACAT 243

QY 1767 CTTGAGATTAACCTGGACACGAGGTTTGACATGGAACAAACATGTTACCGAGCGCT 1826
DB 244 CTTGAGATTAACCTGGACACGAGGTTTGACATGGAACAAACATGTTACCGAGCGCT 303

QY 1827 CTTGGCTCTCCAGCGGGAGTGTCTGCTGTCTGATTAACAGAGTACCGGAGACACCGT 1886
DB 304 CTTGGCTCTCCAGCGGGAGTGTCTGCTGTCTGATTAACAGAGTACCGGAGACACCGT 363

QY 1887 TCTGTGTACTGTGCGCTGCGCAGCTTCGCTGAGCTGACCTATCAGTATCGGCAGAA 1946
DB 364 TCTGTGTACTGTGCGCTGCGCAGCTTCGCTGAGCTGACCTATCAGTATCGGCAGAA 423

QY 1947 CATTCCTGCTTCCGAGTTGCGAGTGGCCGTAACATCCCGTCTCTGAGTGTACTGGGGCG 2006
DB 424 CATTCCTGCTTCCGAGTTGCGAGTGGCCGTAACATCCCGTCTCTGAGTGTACTGGGGCG 483

QY 2007 TAACTGCGGACACTCAGGTGAAGCTCACCAGCGCATGAATTCATATCTGTGAACA 2066
DB 484 TAACTGCGGACACTCAGGTGAAGCTCACCAGCGCATGAATTCATATCTGTGAACA 543

QY 2067 GACAAGGTTCAAAAATTAAGCATCCAGAGCGCTTCAGGAGCTTTTCAGCACTGGAGGTGAA 2126
DB 544 GACAAGGTTCAAAAATTAAGCATCCAGAGCGCTTCAGGAGCTTTTCAGCACTGGAGGTGAA 603

QY 2127 GAGAGCGTGTGTAAAAATACAGAGCAAGACGCTCAAGGTGTTCACAGCGCCCTGAG 2186

DB 604 GAGAGCGTGTGTAAAAATACAGAGCAAGACGCTCAAGGTGTTCACAGCGCCCTGAG 663

QY 2187 GGAAGGAGCGAGGCTCTCCGACAGTGTCTGTGGGTGACTCTTCTGTGAGGTTTTAC 2246

DB 664 GGAAGGAGCGAGGCTCTCCGACAGTGTCTGTGGGTGACTCTTCTGTGAGGTTTTAC 723

QY 2247 CCTCTGAGTGAGACCTCTCCCGAGAGCCCC 2275

DB 724 CCTCTGAGTGAGACCTCTCCCGAGAGCCCC 752

RESULT 5

BP971929

LOCUS

DEFINITION

BP971929

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BP971929 897 bp mRNA linear EST 22-JAN-2001
602240332F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4329079 5',
mRNA sequence.

BP971929
BP971929.1 GI:12339144
EST.

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 897)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LHCMI89 row: p column: 08
High quality sequence stop: 728.

FEATURES

source

1. 897
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:4329079"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 46"

/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 24.2%; Score 647; DB 10; Length 897;

Best Local Similarity 100.0%; Pred. No. 4.7e-120; Indels 0; Gaps 0;

Matches 647; Conservative 0; Mismatches 0;

QY 1908 GCGCAGCTTCGCTGAGTGCACCTATCAGTATCGGAGAACATTCCTGCTCCAGTTGCC 1967

DB 1 GCGCAGCTTCGCTGAGTGCACCTATCAGTATCGGAGAACATTCCTGCTCCAGTTGCC 60

QY 1968 AGTGGCGTAAACATCCCGTCTGACTGCTGCGGCGGTAAGTGGCGACTCAGGTGA 2027

DB 61 AGTGGCGTAAACATCCCGTCTGACTGCTGCGGCGGTAAGTGGCGACTCAGGTGA 120

QY 2028 AGCTCACACGCGCAATGAATTCATCATATCTGTGAACAGACAGGTTCAAAACTAAGC 2087

Db 121 AGCTACACCGCATGAATTCATATCTGTGAACAGACAGGTTCAAAAATAAGC 180
QY 2088 ATCCAGAGCCCTGAGCAGCTTTACAGCACTGGAGGTGAGAGAGCGTGTGTTTAAATATAC 2147
Db 181 ATCCAGAGCCCTGAGCAGCTTTACAGCACTGGAGGTGAGAGAGCGTGTGTTTAAATATAC 240
QY 2148 AGAGCAAGCAGCTCAAGTGTGTTTACAGCCCTGAGGGAAGGAGCGAGGCTCTCG 2207
Db 241 AGAGCAAGCAGCTCAAGTGTGTTTACAGCCCTGAGGGAAGGAGCGAGGCTCTCG 300
QY 2208 ACAGGTGCTCTGGGTGAGCTCTCTGTGAGCTGTTTACCTCTCAGTGAGACCTCCCC 2267
Db 301 ACAGGTGCTCTGGGTGAGCTCTCTGTGAGCTGTTTACCTCTCAGTGAGACCTCCCC 360
QY 2268 AGAGCCCTGGGGCCGACAGCCCTCTCTGTGAGCTGTTTACCTCTCAGTGAGACCTCCCC 2327
Db 361 AGAGCCCTGGGGCCGACAGCCCTCTCTGTGAGCTGTTTACCTCTCAGTGAGACCTCCCC 420
QY 2328 TCAGCAGCAGACGAGCGCTTTCTGTACATGCGGCGCTCCCGCCGAGAGGGGAGTTT 2387
Db 421 TCAGCAGCAGACGAGCGCTTTCTGTACATGCGGCGCTCCCGCCGAGAGGGGAGTTT 480
QY 2388 TGCTCTTTTGTACATTTCCGAACTACAGTTTAAAGCAGAGTCTGTTTTCAGGAAAAGT 2447
Db 481 TGCTCTTTTGTACATTTCCGAACTACAGTTTAAAGCAGAGTCTGTTTTCAGGAAAAGT 540
QY 2448 TTCAAGGGAAGGGCAAGTTTATCAAAAACATGTTTTCAGGAAAGGAGGAGTAAAGTTT 2507
Db 541 TTCAAGGGAAGGGCAAGTTTATCAAAAACATGTTTTCAGGAAAGGAGGAGTAAAGTTT 600
QY 2508 ACAGCTTACAGGACGTACACATATCTGCTGCTGGGAAAACACACAG 2554
Db 601 ACAGCTTACAGGACGTACACATATCTGCTGCTGGGAAAACACACAG 647

RESULT 6
BG395714 840 bp mRNA linear EST 12-MAR-2001
LOCUS 60245835F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4580633 5',
DEFINITION mRNA sequence.

ACCESSION BG395714
VERSION BG395714.1 GI:13289162
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 840)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM300 row: a column: 18
High quality sequence stop: 775.

FEATURES
Location/Qualifiers
1..840
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4580633"
/issue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_16"
/note="Organ: eye; Vector: pOTB7; Site: 1; XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',

adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 23.1%; Score 620; DB 12; Length 840;
Best Local Similarity 99.9%; Pred. No. 1.2e-114; Indels 0; Gaps 0;
Matches 670; Conservative 0; Mismatches 1;
QY 1294 GAAGGGAGTTTCAGAGGACCTGCTGGAGCTGTCAGACGTTTACAGTGTGAGTCTCAGACATT 1353
Db 2 GAAGGGAGTTTCAGAGGACCTGCTGGAGCTGTCAGACGTTTACAGTGTGAGTCTCAGACATT 61
QY 1354 AGCCAGCATTACGTCGTGTGCGGAGCTGTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1413
Db 62 AGCCAGCATTACGTCGTGTGCGGAGCTGTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 121
QY 1414 CCCCACTGCCAGCACCCGAGGCGAGCCAGGAGCCCAAGGCCCCTGGGGAGTGCACCC 1473
Db 122 CCCCACTGCCAGCACCCGAGGCGAGCCAGGAGCCCAAGGCCCCTGGGGAGTGCACCC 181
QY 1474 TCCACGTCGCTCAGCTTCAGCAGCAGTTCAGGATTACGTTGCTGCTCTGCAAGGAGC 1533
Db 182 TCCACGTCGCTCAGCTTCAGCAGCAGTTCAGGATTACGTTGCTGCTCTGCAAGGAGC 241
QY 1534 CAGCCCTGTGCACCTGCTGCTCCAGCCATGCGCCGAGAGCGGAGCGGAGCAG 1593
Db 242 CAGCCCTGTGCACCTGCTGCTCCAGCCATGCGCCGAGAGCGGAGCGGAGCAG 301
QY 1594 GACCCGCTGTGCGCCCTCAGCAGTGTGCGGTGCTGCTGAGCCTTTCTGCCACTGTAC 1653
Db 302 GACCCGCTGTGCGCCCTCAGCAGTGTGCGGTGCTGCTGAGCCTTTCTGCCACTGTAC 361
QY 1654 TGGGGCTGCACCCGAGCGGCTGCTAGCGTGCCTGCCCGCTGCTGAGCTCAACTG 1713
Db 362 TGGGGCTGCACCCGAGCGGCTGCTAGCGTGCCTGCCCGCTGCTGAGCTCAACTG 421
QY 1714 GGTGCAAGTGTCTGGACGCGCTGCTGAAACAAACAGCTACGAGTACAGATCCTGAAG 1773
Db 422 GGTGCAAGTGTCTGGACGCGCTGCTGAAACAAACAGCTACGAGTACAGATCCTGAAG 481
QY 1774 AATTACCTGGCAACAGAGTTTACATGGAACAAACATGTTGACCGAGAGCTCGTGGCT 1833
Db 482 AATTACCTGGCAACAGAGTTTACATGGAACAAACATGTTGACCGAGAGCTCGTGGCT 541
QY 1834 CTCAGAGGGGAGTGTGTTCTGCTGCTGATTACAGAGTACAGGAGACACCGTTCTGTGT 1893
Db 542 CTCAGAGGGGAGTGTGTTCTGCTGCTGATTACAGAGTACAGGAGACACCGTTCTGTGT 601
QY 1894 TACTGCTGTGCGCTGCGCAGCTTCGCTGAGTGTGACCTATCAGTATCGCAGAACTTCCT 1953
Db 602 TACTGCTGTGCGCTGCGCAGCTTCGCTGAGTGTGACCTATCAGTATCGCAGAACTTCCT 661
QY 1954 GCTTCCGAGTT 1964
Db 662 GCTTCCGAGTT 672

RESULT 7
AUI34937 680 bp mRNA linear EST 02-AUG-2002
LOCUS AUI34937 PLACE1 Homo sapiens cDNA clone PLACE1000852 5', mRNA
DEFINITION sequence.
ACCESSION AUI34937
VERSION AUI34937.1 GI:10995476
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)

AUTHORS Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, I., Sugano, S. and Isogai, T.

TITLE HRI human CDNA project

JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human CDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES	source	Location/Qualifiers
		1. .580
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="PLACE100852"
		/tissue_type="placenta"
		/clone_lib="PLACE1"
		/notes="Vector: pME18SFL3"
ORIGIN		
Query Match	22.4%;	Score 600; DB 9; Length 680;
Best Local Similarity	100.0%;	Pred. No. 1.4e-110;
Matches 600; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1503	CGAGATTACGTTGGCCCTTCGAAGGAGCCACGCCCTGTGCACCTGCTGCTTCAGCC 1562
Db	1	CCAGGATTACGTTGGCCCTTCGAAGGAGCCACGCCCTGTGCACCTGCTGCTTCAGCC 60
QY	1563	CATGCCCGACCGGAGAGCGGAGCGGAGCAGGACCCGCGTGTGCCCCCTCAGCAGTGTGC 1622
Db	61	CATGCCCGACCGGAGAGCGGAGCGGAGCAGGACCCGCGTGTGCCCCCTCAGCAGTGTGC 120
QY	1623	GGTCTGCTGCAGCCCTTTCTGCCACTGTACTGGGGCTGCACCCGACCGGCTGCTACGG 1682
Db	121	GGTCTGCTGCAGCCCTTTCTGCCACTGTACTGGGGCTGCACCCGACCGGCTGCTACGG 180
QY	1683	CTGCCTGGCCCCGTTTGTGAGCTCAACCTGGGCTGCAAGTGCTGCACGGCGTGTGAA 1742
Db	181	CTGCCTGGCCCCGTTTGTGAGCTCAACCTGGGCTGCAAGTGCTGCACGGCGTGTGAA 240
QY	1743	CAACACAGCTACGAGTCAGACATCCTTGAGAAATTACCTGGCAACACAGAGTTTGACATG 1802
Db	241	CAACACAGCTACGAGTCAGACATCCTTGAGAAATTACCTGGCAACACAGAGTTTGACATG 300
QY	1803	GAAGAACATGTTGACCGAGAGCCTCGTGGCTCTCCAGCGGGAGTGTTTCTGCTGCTGA 1862
Db	301	GAAGAACATGTTGACCGAGAGCCTCGTGGCTCTCCAGCGGGAGTGTTTCTGCTGCTGA 360
QY	1863	TTACAGAGTCACGGGAGACACCGTTCTGTGTACTGCTGFGGCGTCGGCAGCTTCGGTGA 1922
Db	361	TTACAGAGTCACGGGAGACACCGTTCTGTGTACTGCTGFGGCGTCGGCAGCTTCGGTGA 420
QY	1923	GCTGACCTATCAGTATCGGAGAACATTCCTGCTTCCGAGTTGCGAGTGGCGGTAAACATC 1982
Db	421	GCTGACCTATCAGTATCGGAGAACATTCCTGCTTCCGAGTTGCGAGTGGCGGTAAACATC 480
QY	1983	CCGTCCTGACTGCTACTGGGCGCGTAACTGCGCACTCAGGTGAAAGCTCACCACGCCAT 2042
Db	481	CCGTCCTGACTGCTACTGGGCGCGTAACTGCGCACTCAGGTGAAAGCTCACCACGCCAT 540
QY	2043	GAAATTCAAATCATATCTGTGAACAGACAGGTTCAAAAATTAAGCATCCAGAGCCCTGA 2102
Db	541	GAAATTCAAATCATATCTGTGAACAGACAGGTTCAAAAATTAAGCATCCAGAGCCCTGA 600

RESULT 8

BQ215731
LOCUS
DEFINITION
AGENCOURT_7549119 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059147
5', mRNA sequence.
BQ215731
BQ215731.1 GI:20397131
EST.
Homo sapiens (human)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
1 (bases 1 to 870)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTP/DTP/Gazdar
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
cDNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1325 row: f column: 12
High quality sequence stop: 601.
Location/Qualifiers
1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6059147"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_68"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

ORIGIN	Query Match	21.8%; Score 583; DB 13; Length 870;
	Best Local Similarity 100.0%; Pred. No. 2.7e-107; Indels 0; Gaps 0;	
	Matches 583; Conservative 0; Mismatches 0;	
Qy	916 CAACCGCGTAGAAATGCCAAAACCGTCCACGAGGAGCGTCAGACGACGGCTGGGAAGCCA 975	
Db	1 CAACCGCGTAGAAATGCCAAAACCGTCCACGAGGAGCGTCAGACGACGGCTGGGAAGCCA 60	
Qy	976 GACAAGATGGAGAGACGCTGACATGATCATCTGCCAGAACCTGTCTGCAGACTGGGTG 1035	
Db	61 GACAAGATGGAGAGACGCTGACATGATCATCTGCCAGAACCTGTCTGCAGACTGGGTG 120	
Qy	1036 AGTTTCGACGCCCTGCATGCACAGCTTCTGCGCGGCTTGCTACTCGGCGCTCGATGGAGCGC 1095	
Db	121 AGTTTGCAGCCCTGCGATGCACAGCTTCTGCGCGGCTTGCTACTCGGCGCTCGATGGAGCGC 180	
Qy	1096 TCGTCCCGTGTGCTCTACCTCGCGCTGTCCCGTGGAGCGGATCTGTAAAAACCACTCTCTC 1155	
Db	181 TCGTCCCGTGTGCTCTACCTCGCGCTGTCCCGTGGAGCGGATCTGTAAAAACCACTCTCTC 240	
Qy	1156 AACAACTCTGTGAAGCATACCTCATCCAGCATCCAGACAAGAGTCCGAGTGAAGAAGAT 1215	
Db	241 AACAACTCTGTGAAGCATACCTCATCCAGCATCCAGACAAGAGTCCGAGTGAAGAAGAT 300	
Qy	1216 GTGCAAAGTATGGATGCCAGGAATAAAATCACTCAAGACATGCTGCAGCCCCAAAGTCAGG 1275	
Db	301 GTGCNAAAGTATGGATGCCAGGAATAAAATCACTCAAGACATGCTGCAGCCCCAAAGTCAGG 360	
Qy	1276 CGGTCCTTTTCTGATGAAGAAGGAGTTTCAGAGGACCTGCTGGAGCTGTTCAGACGTTGAC 1335	
Db	361 CGGTCCTTTTCTGATGAAGAAGGAGTTTCAGAGGACCTGCTGGAGCTGTTCAGACGTTGAC 420	

QY 1336 AGTGAGTCCTCAGACATTAGCCAGCCATAGCTCGTGTGGCGGAGTGCTCTGAGTACAGA 1395
 Db 421 AGTGAGTCCTCAGACATTAGCCAGCCATAGCTCGTGTGGCGGAGTGCTCTGAGTACAGA 480
 QY 1396 AGGCAGGCGGCGAGCCTCCCTCCACTGCCAGCACCAGGCGGAGCAGGAGCCCCACAG 1455
 Db 481 AGGCAGGCGGCGAGCCTCCCTCCACTGCCAGCACCAGGCGGAGCAGGAGCCCCACAG 540
 QY 1456 GCCTGTGGGGATGACACCTCCAGCTCGCTCAGCTGACGACAG 1498
 Db 541 GCCTGTGGGGATGACACCTCCAGCTCGCTCAGCTGACGACAG 583

RESULT 9
 BM459647
 LOCUS
 DEFINITION AGENCOURT_6418058 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534277
 5', mRNA sequence.
 ACCESSION BM459647
 VERSION BM459647.1 GI:18508687
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 916)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM12220 row: h column: 22
 High quality sequence stop: 642.
 Location/Qualifiers
 1..916
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5534277"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb."

FEATURES
 source
 Query Match 21.1%; Score 565; DB 12; Length 916;
 Best Local Similarity 100.0%; Pred. No. 9.7e-104; Mismatches 0; Indels 0; Gaps 0;
 Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 709 GGGTCTGGGGGTGGTGGCATCTCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTATGAA 768
 Db 155 GGGTCTGGGGGTGGTGGCATCTCCCTAAAGGAAGTGGTCCCTCTGTGGCAAGTATGAA 214
 QY 769 GTCTCCAGCTTTGGCTCAGCTCTCCCAAGAAAGACTGGCTCTTTTGTGGTGGAA 828
 Db 215 GTCTCCAGCTTTGGCTCAGCTCTCCCAAGAAAGACTGGCTCTTTTGTGGTGGAA 274
 QY 829 CCCAGGATCAGGAGGATTTGGAGCCCTGAGAGAAATGAGAGGATGGGACCTT 888
 Db 275 CCCAGGATCAGGAGGATTTGGAGCCCTGAGAGAAATGAGAGGATGGGACCTT 334
 QY 889 GACCTGAACGGGCGAGTTGGTGGCGCAACCCGGGTAGAAATGCCCAACCGTCCAGAG 948
 Db 335 GACCTGAACGGGCGAGTTGGTGGCGCAACCCGGGTAGAAATGCCCAACCGTCCAGAG 394

QY 949 GACGTACAGCAGCGGCTGGGAAGCCAGACAGATGAGGAGCGCTGACATGCATCATC 1008
 Db 395 GACGTACAGCAGCGGCTGGGAAGCCAGACAGATGAGGAGCGCTGACATGCATCATC 454
 QY 1009 TGCCAGGACCTGCTGCACGACTGGGTGAGTTTGAGCCCTGCATGCACACGTTCTGCGCG 1068
 Db 455 TGCCAGGACCTGCTGCACGACTGGGTGAGTTTGAGCCCTGCATGCACACGTTCTGCGCG 514
 QY 1069 GCTTGCTACTCGGGCTGGATGGAGCGCTGCTCCTCTGTGCTTACTGCGCTGTCCTG 1128
 Db 515 GCTTGCTACTCGGGCTGGATGGAGCGCTGCTCCTCTGTGCTTACTGCGCTGTCCTG 574
 QY 1129 GAGCGGATCTGTAAGAACCCATCTCAACACCTCGTGGAGCATATCTCATCCAGCAT 1188
 Db 575 GAGCGGATCTGTAAGAACCCATCTCAACACCTCGTGGAGCATATCTCATCCAGCAT 634
 QY 1189 CCAGACAAGAGTCCGAGTGAAGAAGATGTCAAAAGTATGAGATGCCAGAAATAATCACT 1248
 Db 635 CCAGACAAGAGTCCGAGTGAAGAAGATGTCAAAAGTATGAGATGCCAGAAATAATCACT 694
 QY 1249 CAAGACATGCTGCAGCCCAAGATCA 1273
 Db 695 CAAGACATGCTGCAGCCCAAGATCA 719

RESULT 10
 BM542545
 LOCUS
 DEFINITION AGENCOURT_6426028 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5520909
 5', mRNA sequence.
 ACCESSION BM542545
 VERSION BM542545.1 GI:18772149
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1143)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM12185 row: k column: 22
 High quality sequence stop: 666.
 Location/Qualifiers
 1..1143
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5520909"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb."

FEATURES
 source
 Query Match 21.0%; Score 563; DB 12; Length 1143;
 Best Local Similarity 100.0%; Pred. No. 2e-103; Mismatches 0; Indels 0; Gaps 0;
 Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
 QY 2099 CTGACGAGCTTTCAGCACTGGAGGTGAAGAGCGCTGTTTAAATACAGAGACAAGCA 2158

RESULT	11
LOCUS	BE407732
DEFINITION	60129977F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3629682 5', linear mRNA 666 bp EST 21-JUN-2000
ACCESSION	B5407732
VERSION	B5407732
KEYWORDS	mRNA sequence.
SOURCE	B5407732.1 GI:9344182
ORGANISM	Homo sapiens (human)

```

FEATURES
source
Location/Qualifiers
1..666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3629682"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"

```

ORIGIN

Query Match	20.9%	Score 560;	DB 10;	Length 666;
Best Local Similarity	99.8%	Fred. No. 1.38-102;		
Matches 610;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps 0;				

RESULT 12	AL539797/c	1201 bp	mRNA	linear	EST 31-MAY-2003
LOCUS	AL539797				
DEFINITION	AL539797 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone				
	CS00F031YA23 3-PRIME, mRNA sequence.				
ACCESSION	AL539797				
VERSION	AL539797.2	GI:31264360			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1201)				

```
AUTHORS      Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      On Feb 15, 2001 this sequence version replaced gi:12869347.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6792.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF031AA12NP1&cluster=6792.r. Contact :
Peng Liang Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF031AA12NP1.
FEATURES
    source
        1..1201
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DF031YA23"
            /tissue_type="FETAL BRAIN"
            /dev_stage="fetal"
            /clone_lib="Homo sapiens FETAL BRAIN"
            /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
            was primed with a NotI-oligo(dT) primer. Five prime end
            enriched, double-strand cDNA was digested with Not I and
            cloned into the Not I and EcoRV sites of the pCMVSPORT 6
            vector. Library was not normalized."
ORIGIN
    Query Match      20.1%; Score 539; DB 9; Length 1201;
    Best Local Similarity 100.0%; Pred. No. 1.1e-98;
    Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1704 GCTCAACCTGGGTGACAAGTGTGTGACGGCGTGTGAACACACAGCTACAGTCA 1763
DB 940 GCTCAACCTGGGTGACAAGTGTGTGACGGCGTGTGAACACACAGCTACAGTCA 881
QY 1764 CATCTGAGAAATTAAGTGTGACCAACAGAGTTTGACATGGAATAATGTCACCGAGAG 1823
DB 880 CATCTGAGAAATTAAGTGTGACCAACAGAGTTTGACATGGAATAATGTCACCGAGAG 821
QY 1824 CTTCTGGTCTCCAGCGGGAGTGTTCGTGTCTGATTAACAGAGTCAACGGAGACAC 1883
DB 820 CTTCTGGTCTCCAGCGGGAGTGTTCGTGTCTGATTAACAGAGTCAACGGAGACAC 761
QY 1884 CTTCTGTGTACTGTGGCTGGCGAGCTTCCGTGAGCTGACCTATCAGTATCGGCA 1943
DB 760 CTTCTGTGTACTGTGGCTGGCGAGCTTCCGTGAGCTGACCTATCAGTATCGGCA 701
QY 1944 GAACATTCCTGCTCCAGGTGCGAGTGGCGGTAACATCCCGTCTGACTGTCTGCGG 2003
DB 700 GAACATTCCTGCTCCAGGTGCGAGTGGCGGTAACATCCCGTCTGACTGTCTGCGG 641
QY 2004 CCGTAAGTCCGCACTCAGGTGAAGTCAACAGCGCATGAATTCATATCTGTGA 2063
DB 640 CCGTAAGTCCGCACTCAGGTGAAGTCAACAGCGCATGAATTCATATCTGTGA 581
QY 2064 ACAGAGAGGTTTCAAAACTAAGCATCCAGAGCCCTGAGCAGCTTTTCAGCACTGAGGT 2123
DB 580 ACAGAGAGGTTTCAAAACTAAGCATCCAGAGCCCTGAGCAGCTTTTCAGCACTGAGGT 521
QY 2124 GAAGAGAGCGTGTGTTTTAAATAACAGAGCAACAGCTCAAGGTGTTTTACAGCCCTCT 2183
DB 520 GAAGAGAGCGTGTGTTTTAAATAACAGAGCAACAGCTCAAGGTGTTTTACAGCCCTCT 461
QY 2184 GAGGAGAGGACCGAGGTTCTCGACAGGTGCTCTGGGTGACTCTTCTGTGAGCTTT 2242
DB 460 GAGGAGAGGACCGAGGTTCTCGACAGGTGCTCTGGGTGACTCTTCTGTGAGCTTT 402
```

```
RESULT 13
BF971838
LOCUS      602240424f1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4329076 5',
DEFINITION mRNA sequence.
ACCESSION BF971838
VERSION    BF971838.1 GI:12339053
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 934)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgsb@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM189 row: p column: 05
High quality sequence stop: 748.
FEATURES
    source
        1..934
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="IMAGE:4329076"
            /tissue_type="leiomyosarcoma cell line"
            /lab_hosts="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_46"
            /notes="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCAACGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)
            using ZAP-cDNA synthesis kit (Stratagene) and Superscript
            II RT (Life Technologies). Note: this is a NIH_MGC
            Library."
ORIGIN
    Query Match      19.4%; Score 519; DB 10; Length 934;
    Best Local Similarity 99.8%; Pred. No. 1.4e-94;
    Matches 539; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1908 GCGCAGCTTCGTCGAGTACCTATCAGTATCGGCAGACATTCCTGCTCCGAGTTGCC 1967
DB 1 GCGCAGCTTCGTCGAGTACCTATCAGTATCGGCAGACATTCCTGCTCCGAGTTGCC 60
QY 1968 AGTGGCCGTAACATCCCGTCTGCTGCTACTGGGGCGGTAACCTGCGCACTCAGGTGAA 2027
DB 61 AGTGGCCGTAACATCCCGTCTGCTGCTACTGGGGCGGTAACCTGCGCACTCAGGTGAA 120
QY 2028 AGCTCACACGCCATGAATTCATATCTGTGAACAGCAAGGTTCAAAAATAAGC 2087
DB 121 AGCTCACACGCCATGAATTCATATCTGTGAACAGCAAGGTTCAAAAATAAGC 180
QY 2088 ATCCAGAGCCCTGAGCAGCTTTCAGGACCTGGAGAGCGTGTGTTTTAAATAATAC 2147
DB 181 ATCCAGAGCCCTGAGCAGCTTTCAGGACCTGGAGAGCGTGTGTTTTAAATAATAC 240
QY 2148 AGAGCAAGCAGTCAAGGTGTTTTTCACGCCCTGAGGGAAGGAGCGAGGCTCTCCG 2207
DB 241 AGAGCAAGCAGTCAAGGTGTTTTTCACGCCCTGAGGGAAGGAGCGAGGCTCTCCG 300
QY 2208 ACAGGTGCTGGGGTGACTCTCTGTGGAGCTTTTACCTCTGAGTGACACCTCCCC 2267
DB 2208 ACAGGTGCTGGGGTGACTCTCTGTGGAGCTTTTACCTCTGAGTGACACCTCCCC 2267
```


Db 301 ACAGTGCTCTGGGTGACTCTTCTGTGGAGCTTTTACCTCTGTAGTGAGACCTCCCC 360

QY 2268 AGAGCCCCGGGGCGGAGCCGCCCTCTCTGGTGAAGCGCTGGGCGAGGGCTCGTGTGGCA 2327

Db 361 AGAGCCCCGGGGCGGAGCCCG-CCCTCCTGGTGAAGCGCTGGGCGAGGGCTCGTGTGGCA 419

QY 2328 TCAGCAGCAGAGACGAGCCCTTCTGTAAACATGCGGGCGCTCCCGCCGAGAGGGCGAGTTT 2387

Db 420 TCAGCAGCAGAGACGAGCCCTTCTGTAAACATGCGGGCGCTCCCGCCGAGAGGGCGAGTTT 479

QY 2388 TGCTCTTTTGTACATTTTCCGAACATACAGTTAAAGCAGAAGTCTGTTCAGGAAAGT 2447

Db 480 TGCTCTTTTGTACATTTTCCGAACATACAGTTAAAGCAGAAGTCTGTTCAGGAAAGT 539

QY 2448 TTCAGGGAGAGGCGCAAGTTTATCAAAACATTTTTCAGGAGAGGGAGCATAAGTTT 2507

Db 540 TTCAGGGAGAGGCGCAAGTTTATCAAAACATTTTTCAGGAGAGGGAGCATAAGTTT 599

QY 2508 ACAGCCTACAGAGCTACACAAATATCTCTGCTGTGGGAAA 2547

Db 600 ACAGCCTACAGAGCTACACAAATATCTCTGCTGTGGGAAA 639

RESULT 14

CA916737

LOCUS

DEFINITION 729 bp mRNA linear EST 14-MAR-2003

30LSPRIM2-5A human lymphocyte Matchmaker cDNA library Homo sapiens cDNA clone L5 5' similar to CHFR RF; cell cycle check point, mRNA sequence.

ACCESSION CA916737

VERSION CA916737.1 GI:28951974

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 729)

TITLE Scarafia L.E., Stouffer S.D. and Swinney D.C.

JOURNAL Identification of Ring Finger proteins that interact with Ubch5a, an ubiquitin-conjugating enzyme

COMMENT unpublished (2002)

Contact: Scarafia LE

Enzymology

Roche Bioscience

3401 Hillview Ave, S3-1, Palo Alto, CA 94304-1397, USA

Tel: 650 354 7997

Fax: 650 354 7554

Email: liliana.scarafia@roche.com

Seq primer: matchmaker 5' AD (ctattcgatgatgaagatccccacccaccc)

POLYA=No.

FEATURES

source

1. 729

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="L5"

/cell_type="B cell"

/cell_lines="EBV-transformed human peripheral blood lymphocyte; B cell population; IG (+)"

/lab_host="yeast/2.coli"

/clone_lib="human lymphocyte Matchmaker cDNA library"

/note="Vector: pACT; Site 1: XhoI; Site 2: XhoI; Matchmaker yeast two-hybrid system from Clontech; pACT has GAL4 AD under ADHI promoter. This oligo-dr primed library was screened with human Ubch5a as bait, to obtain interacting proteins."

ORIGIN

Query Match 19.1%; Score 513; DB 14; Length 729;

Best Local Similarity 99.5%; Pred. No. 2.6e-93;

Matches 663; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 560 CCGATCCCGGGTCCCTCCGTGTCGCCGCCACTCAGTGTGCTTTGAGGAACACAGC 619

Db 13 CCGATCCCGGGTCCCTCCGTGTCGCCGCCACTCAGTGTGCTTTGAGGAACACAGC 72

QY 620 CATCAACATCAGCTCAGACCTCTTCCCAACAGCTCCGCCCTCTTCCACGAGAGCCCTTC 679

Db 73 CATCAACATCAGCTCAGACCTCTTCCCAACAGCTCCGCCCTCTTCCACGAGAGCCCTTC 132

QY 680 CTGCAAGGGCGAGAGCGTTCTCCAGTTGTGGGTCTGCGGGGTGGTGGCATCTCCCTAAAG 739

Db 133 CTGCAAGGGCGAGAGCGTTCTCCAGTTGTGGGTCTGCGGGGTGGTGGCATCTCCCTAAAG 192

QY 740 GAATGCTCCCTCTGTGGCAAGTGAAGTCTCCAGCTTTGCCCTCAGCTCTCCAGACA 799

Db 193 GAATGCTCCCTCTGTGGCAAGTGAAGTCTCCAGCTTTGCCCTCAGCTCTCCAGACA 252

QY 800 GAAAGACTCGCTCCCTTTTCGTGCTTGGAAACCCAGGATCAGAGGATTGGAGCCCGTGA 859

Db 253 GAAAGACTCGCTCCCTTTTCGTGCTTGGAAACCCAGGATCAGAGGATTGGAGCCCGTGA 312

QY 860 AGAAGAAATGAGAGAGATGGGACCTTGACCTGAACGGGAGTGTGTGCGCACAC 919

Db 313 AGAAGAAATGAGAGAGATGGGACCTTGACCTGAACGGGAGTGTGTGCGCACAC 372

QY 920 CCGGTAGAAATGCCAAACCGTCCACAGGAGCTCAGAGCAGCGCTGGGAAGCCAGACA 979

Db 373 CCGGTAGAAATGCCAAACCGTCCACAGGAGCTCAGAGCAGCGCTGGGAAGCCAGACA 432

QY 980 AGATGAGAGAGAGCTGACATCATCATCTGCGAGGACTGCTGCGAGACTCGCGTAGTT 1039

Db 433 AGATGAGAGAGAGCTGACATCATCATCTGCGAGGACTGCTGCGAGACTCGCGTAGTT 492

QY 1040 TGCAGCCCTGTCATGCACAGCTTCTCGGGGCTTGTCTCGGGCTGGATGGAGCGCTCGT 1099

Db 493 TGCAGCCCTGTCATGCACAGCTTCTCGGGGCTTGTCTCGGGCTGGATGGAGCGCTCGT 552

QY 1100 CCTGTGCTCTACCTCCCGCTCTCCCGTGGAGCGGATCTGTAAAAACCAATCTCAACA 1159

Db 553 CCTGTGCTCTACCTCCCGCTCTCCCGTGGAGCGGATCTGTAAAAACCAATCTCAACA 612

QY 1160 ACTCTGTGAAGCATACCTCATCCAGCATCCAGAGCTCGAGTGAAGAGATGTGC 1219

Db 613 ACTCTGTGAAGCATACCTCATCCAGCATCCAGAGCTCGAGTGAAGAGATGTGC 672

1220 AAAGTA 1225

673 AAAGTA 678

RESULT 15

BE280165

LOCUS

DEFINITION 690 bp mRNA linear EST 13-JUL-2000

601158370F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3504928 5', mRNA sequence.

ACCESSION BE280165

VERSION BE280165.1 GI:9155072

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 690)

NIH-MGC <http://mgi.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM180 row: 1 column: 17

High quality sequence stop: 626.
Location/Qualifiers
1. .690
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3504928"
/tissue type="choriocarcinoma"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH MGC 21"
/notes="Organ: Placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 18.2%; Score 488; DB 10; Length 690;
Best Local Similarity 99.8%; Pred. No. 2.6e-88;
Matches 538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1156 AACAACTCGTGAAGCATACCTCACCAGCATCCAGACAGAGTCGCGAGTGAAGAGAT 1215
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 AACAACTCGTGAAGCATACCTCACCAGCATCCAGACAGAGTCGCGAGTGAAGAGAT 60
QY 1216 GTGCAAGTATGATGCCAGGATAAATCACTCAAGACATGCTGCAGCCCAAGTCAGG 1275
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 GTGCAAGTATGATGCCAGGATAAATCACTCAAGACATGCTGCAGCCCAAGTCAGG 120
QY 1276 CGGTCTTTTCTGATGAAGAGGAGTTTCAGAGGACCTCTGCAGAGTGTTCAGACGTTGAC 1335
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 CGGTCTTTTCTGATGAAGAGGAGTTTCAGAGGACCTCTGCAGAGTGTTCAGACGTTGAC 180
QY 1336 AGTGAGTCTCTCAGACATTAGCCAGCCATACCTCGTGTGCCGCGAGTGTCTGAGTACAGA 1395
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 AGTGAGTCTCTCAGACATTAGCCAGCCATACCTCGTGTGCCGCGAGTGTCTGAGTACAGA 240
QY 1396 AGGCAGCGCGGAGCCTCCCACTGCCCCAGCACCCGAGGCGGAGCCAGGCCCCACAG 1455
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 AGGCAGCGCGGAGCCTCCCACTGCCCCAGCACCCGAGGCGGAGCCAGGCCCCACAG 300
QY 1456 GCCTGGGGGATGCACCTCCAGCTCCGTCAGCAGTGAAGACAGTCCAGGATTACGTG 1515
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 GCCTGGGGGATGCACCTCCAGCTCCGTCAGCAGTGAAGACAGTCCAGGATTACGTG 360
QY 1516 TGCCCTCTCAAGGAAGCACGCCCTGTGCACCTGTCTTCCAGCCCATGCCGACCGG 1575
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 TGCCCTCTCAAGGAAGCACGCCCTGTGCACCTGTCTTCCAGCCCATGCCGACCGG 420
QY 1576 AGAGCGGAGCGGAGGAGCCGGGTGTGCGGCTCAGAGTGTGCGGTCTGCGTGCAG 1635
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 AGAGTGGAGCGGAGGAGCCGGGTGTGCGGCTCAGAGTGTGCGGTCTGCGTGCAG 480
QY 1636 CCTTTCTGCACCTGTACTGGGGCTGCACCCGAGCCGGGTGTACGGGTGCTGGCCCC 1694
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 CCTTTCTGCACCTGTACTGGGGCTGCACCCGAGCCGGGTGTACGGGTGCTGGCCCC 539

Search completed: May 15, 2004, 07:58:44
Job time : 10042 secs

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:36:57 ; Search time 136.025 Seconds
(without alignments)
1379.240 Million cell updates/sec

Title: US-10-048-046-2

Perfect score: 3585
Sequence: 1 MERPEKGQPPPPQWGRLL.....VKAHAKFNHCQTRFKN 664

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1380s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3585	100.0	664	4 AAB3843	Aab3843 Amino aci
2	3585	100.0	664	4 AAB20219	Aab20219 Human Chf
3	3504	97.7	652	4 AAB93168	Aab93168 Human pro
4	3168.5	88.4	623	5 AAB97233	Abb97233 Novel hum
5	3162.5	88.2	623	4 AAB93182	Aab93182 Human pro
6	1954.5	54.5	426	4 AAU15856	Aau15856 Human nov
7	1954.5	54.5	426	6 AAU54925	Aau54925 Human nov
8	1364	38.0	269	4 AAO08972	Aao08972 Human pol
9	843.5	23.5	230	4 AAU16317	Aau16317 Human nov
10	843.5	23.5	230	6 ABUS5386	Abu55386 Human nov
11	566	15.8	128	6 ADA54231	Ada54231 Human pro
12	474	13.2	92	4 AAU16156	Aau16156 Human nov
13	474	13.2	92	6 ABUS5225	Abu55225 Human nov
14	432	12.1	99	4 AAU16157	Aau16157 Human nov
15	432	12.1	99	4 AAU16576	Aau16576 Human nov
16	432	12.1	99	6 ABUS5645	Abu55645 Human nov
17	432	12.1	99	6 ABUS5226	Abu55226 Human nov
18	355	9.9	426	3 AAG20353	Aag20353 Arabidops
19	355	9.9	426	3 AAG37659	Aag37659 Arabidops
20	355	9.9	453	3 AAG20352	Aag20352 Arabidops
21	355	9.9	453	3 AAG37658	Aag37658 Arabidops
22	354	9.9	350	3 AAG20354	Aag20354 Arabidops
23	354	9.9	350	3 AAG37660	Aag37660 Arabidops
24	203	5.7	485	4 AAB95017	Aab95017 Human pro
25	160.5	4.5	352	5 AAU80358	Aau80358 Human cel

ALIGNMENTS

RESULT 1

AAB83843
ID AAB83843 standard; protein; 664 AA.

XX AC AAB83843;
XX 22-AUG-2001 (first entry)

DT Amino acid sequence of a human ring finger protein designated FHAR1.

DE FHAR1; RING finger protein; cancer; vaccine.

XX Homo sapiens.

XX WO200142430-A1.

XX 14-JUN-2001.

XX 07-DEC-2000; 2000WO-US033094.

XX 08-DEC-1999; 99US-00456876.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Zhou B, Zhu Y, Chaturvedi P, Hurle MR, Li X;

XX WPI: 2001-381663/40.

XX N-PSDB; AAF89709.

XX New FHAR1 polypeptide, a member of the RING finger protein family for

XX diagnosing and treating cancer, and for use in anti-cancer vaccines.

XX Claim 1; Page 19; 28pp; English.

XX The present sequence represents a FHAR1 polypeptide, which is a member of

XX the RING finger protein family. FHAR1 is useful in the treatment of

XX cancer, and as a vaccine for inducing an immunological response in a

XX mammal. FHAR1 polynucleotides may also be used as a diagnostic reagent

XX through detection of mutations in the associated gene, and for chromosome

XX localization studies, and tissue expression studies. FHAR1 antibodies are

XX useful to isolate and to identify clones expressing the polypeptides, or

XX to purify the polypeptides by affinity chromatography and to treat cancer

XX Sequence 664 AA;

XX Query Match 100.0%; Score 3585; DB 4; Length 664;

XX Best Local Similarity 100.0%; Pred. No. 9.8e-270;

XX Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERPEEGKQSPPPQPPWGRLLRLGAEAGEPHVLLRKREWTIGRRGCDLSFPNKLVSQDH 60
 Db 1 MERPEEGKQSPPPQPPWGRLLRLGAEAGEPHVLLRKREWTIGRRGCDLSFPNKLVSQDH 60
 QY 61 CRIVDEKSGQVLTSTSTGTVINKLVKKQKQTPLOTGVDVILYVRKNEPERNVAYLY 120
 Db 61 CRIVDEKSGQVLTSTSTGTVINKLVKKQKQTPLOTGVDVILYVRKNEPERNVAYLY 120
 QY 121 ESLSEKQGMTOESFEANKENVPHGTGKTSAGAGRGADPRVPPSPATQVCFEFPQST 180
 Db 121 ESLSEKQGMTOESFEANKENVPHGTGKTSAGAGRGADPRVPPSPATQVCFEFPQST 180
 QY 181 TSDLFTTASASTEPSPAGRRSSSCGSGGGISPKGSGPSVASDEVSSFASALPDRKTA 240
 Db 181 TSDLFTTASASTEPSPAGRRSSSCGSGGGISPKGSGPSVASDEVSSFASALPDRKTA 240
 QY 241 SFSSLEFPQDELEPVKKQKRGDGLDLNGQLLVAAQPRNAQTVDHVRRAAGKPDKWE 300
 Db 241 SFSSLEFPQDELEPVKKQKRGDGLDLNGQLLVAAQPRNAQTVDHVRRAAGKPDKWE 300
 QY 301 TLTCTICQDLHDCVSLQPCMTFCAACVSGMERSLLCPTCRCPVERICKNHILNIVE 360
 Db 301 TLTCTICQDLHDCVSLQPCMTFCAACVSGMERSLLCPTCRCPVERICKNHILNIVE 360
 QY 361 AYLIQHDPKSRSEEDVQSMDSANKITQDMLQPKVRSFSDSEGSSEDLLELSDVDS 420
 Db 361 AYLIQHDPKSRSEEDVQSMDSANKITQDMLQPKVRSFSDSEGSSEDLLELSDVDS 420
 QY 421 ISQPVVVCQCEYRQAAOPHCPAPGEPGAPGALGADPSTSVLTAVQDVVCPLOG 480
 Db 421 ISQPVVVCQCEYRQAAOPHCPAPGEPGAPGALGADPSTSVLTAVQDVVCPLOG 480
 QY 481 SHALCTCCFPQPPDRAREQDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAPFCELN 540
 Db 481 SHALCTCCFPQPPDRAREQDPRVAPQCAVCLQPFCHLYWGCTRTGCGCLAPFCELN 540
 QY 541 LGDKCLDGLVNNNSVESDILKYLATRLTWKMLTSLVALQKGVFLSDYRTGDTVL 600
 Db 541 LGDKCLDGLVNNNSVESDILKYLATRLTWKMLTSLVALQKGVFLSDYRTGDTVL 600
 QY 601 CYCCGLRSFRELTYQYRQNPASELPVATSRPDCYMGRCRTQVKAHAMKFNHICEQT 660
 Db 601 CYCCGLRSFRELTYQYRQNPASELPVATSRPDCYMGRCRTQVKAHAMKFNHICEQT 660
 QY 661 RPKN 664
 Db 661 RPKN 664

RESULT 2
 AAB20219
 ID AAB20219 standard; protein; 664 AA.
 XX AC AAB20219;
 XX DT 14-MAY-2001 (first entry)
 XX DE Human Chfr (checkpoint with FHA and ring finger) protein.
 KW Checkpoint with forkhead associated domain and ring finger; Chfr; human;
 KW mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening;
 KW ubiquitin-protein ligase.
 XX OS Homo sapiens.
 XX FH Key
 FT 31..103 Location/Qualifiers
 FT /label= Forkhead-associated_domain
 FT 303..346
 FT /label= Ring_finger-domain
 FT 476..641
 FT /note= "cysteine-rich region"

FT Misc-difference 580 /note= "Met in U2OS cells"
 XX FT
 PN WC200109150-A2.
 XX 08-FEB-2001.
 XX 14-JUN-2000; 2000WO-US016391.
 XX 29-JUL-1999; 99US-0146194P.
 XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
 PA Halazonetis T, Scolnick D;
 XX WPI; 2001-182927/18.
 XX N-PSDB; AAF30352.
 XX Novel nucleic acid sequence of mitotic checkpoint gene encoding a
 PT checkpoint with forkhead-associated domain and ring finger protein, for
 PT diagnosing tumorigenic cells and in screening for anticancer drugs.
 PS Claim 8(a); Fig 4A-C; 85pp; English.
 CC The present sequence is that of human mitotic checkpoint protein Chfr,
 CC having a forkhead associated domain (FHA) and a ring finger domain. The
 CC protein is required for regulation of the transition of cells from
 CC prophase to metaphase during mitosis. It has ubiquitin-protein ligase
 CC activity. The Chfr checkpoint was evident in primary human cells, but was
 CC inactivated in 4 of 8 human cancer cell lines. In U2OS cells, a mutation
 CC was identified that caused a Val to Met amino acid substitution in the
 CC highly conserved C-terminal Cys-rich region of the Chfr protein. In the
 CC absence of the Chfr checkpoint, cells subjected to mitotic stress
 CC condensed their chromosomes despite failing to separate their
 CC chromosomes. Chfr may monitor centrosome separation. Inactivation of the
 CC Chfr gene (see AAF30352) in human cancer is theorized to underlie the
 CC increased sensitivity of cancer cells to antimitotic drugs. Polypeptides
 CC comprising the present sequence, or sequences comprising at least amino
 CC acids 31-103, 303-346 and/or 476-641 of this sequence, are claimed.
 CC Claimed methods of determining the tumorigenic potential of a cell
 CC comprise examining the cell for the presence of Chfr expression or for
 CC the presence of Chfr-mediated ubiquitin-protein ligase activity (in both
 CC cases, absence of expression indicating predisposition to tumourgenesis
 CC upon exposure to mitotic stress). A diagnostic kit for detecting the
 CC tumorigenic potential of cell cells comprises may comprise a ligand that
 CC binds to Chfr, such as an antibody or its fragment. Inhibitors of Chfr
 CC are identified by monitoring their effect on Chfr expression, and are
 CC used to retard the growth of cancer cells
 XX SQ Sequence 664 AA;
 Query Match Similarity 100.0%; Score 3585; DS 4; Length 664;
 Best Local Similarity 100.0%; Pred. No. 9.8e-270;
 Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MERPEEGKQSPPPQPPWGRLLRLGAEAGEPHVLLRKREWTIGRRGCDLSFPNKLVSQDH 60
 Db 1 MERPEEGKQSPPPQPPWGRLLRLGAEAGEPHVLLRKREWTIGRRGCDLSFPNKLVSQDH 60
 QY 61 CRIVDEKSGQVLTSTSTGTVINKLVKKQKQTPLOTGVDVILYVRKNEPERNVAYLY 120
 Db 61 CRIVDEKSGQVLTSTSTGTVINKLVKKQKQTPLOTGVDVILYVRKNEPERNVAYLY 120
 QY 121 ESLSEKQGMTOESFEANKENVPHGTGKTSAGAGRGADPRVPPSPATQVCFEFPQST 180
 Db 121 ESLSEKQGMTOESFEANKENVPHGTGKTSAGAGRGADPRVPPSPATQVCFEFPQST 180
 QY 181 TSDLFTTASASTEPSPAGRRSSSCGSGGGISPKGSGPSVASDEVSSFASALPDRKTA 240
 Db 181 TSDLFTTASASTEPSPAGRRSSSCGSGGGISPKGSGPSVASDEVSSFASALPDRKTA 240
 QY 241 SFSSLEFPQDELEPVKKQKRGDGLDLNGQLLVAAQPRNAQTVDHVRRAAGKPDKWE 300

Db 241 SFSSLEPQDELDLEPVKKKRGDGLDLNGQLLVAQPRNAQTVDHVRAAAGKPKMEE 300
Qy 301 TLTCTICQDILLHDCVSLQPCMTFCAACYSGWMSRSLCPTCRCPVERICKNHLNVLVE 360
Db 301 TLTCTICQDILLHDCVSLQPCMTFCAACYSGWMSRSLCPTCRCPVERICKNHLNVLVE 360
Qy 361 AYLIQHPDKSRSEEDVQSDMARNKITQDMLQPKVRSFDESGSDDLLELSDVDSESD 420
Db 361 AYLIQHPDKSRSEEDVQSDMARNKITQDMLQPKVRSFDESGSDDLLELSDVDSESD 420
Qy 421 ISQPVVVCRCPEYRQAAPPHCPAPEGEPGAQALGDAPSTVSLTAVQDYVCPLOQ 480
Db 421 ISQPVVVCRCPEYRQAAPPHCPAPEGEPGAQALGDAPSTVSLTAVQDYVCPLOQ 480
Qy 481 SHALCTCCFQPMDDRAERQDPRAVAPQCAVCLQPFCHLYWGCTRTGCGCLAPFCELN 540
Db 481 SHALCTCCFQPMDDRAERQDPRAVAPQCAVCLQPFCHLYWGCTRTGCGCLAPFCELN 540
Qy 541 LGDKCLDGLVNNNSYEDILKNYLATRGLTWKNNLTESLVALQRGVFLSDYRVGTDL 600
Db 541 LGDKCLDGLVNNNSYEDILKNYLATRGLTWKNNLTESLVALQRGVFLSDYRVGTDL 600
Qy 601 CYCCGLRSFRELTYQYRQNIIPASELPVAVTSRPDCYWGRCRTQVKAHMKFNHICEQT 660
Db 601 CYCCGLRSFRELTYQYRQNIIPASELPVAVTSRPDCYWGRCRTQVKAHMKFNHICEQT 660
Qy 661 RPKN 664
Db 661 RPKN 664

RESULT 3
ID AAB93168
AA B93168 standard; protein; 652 AA.

XX AAB93168;
XX 26-JUN-2001 (first entry)
XX Human protein sequence SEQ ID NO:12100.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.

XX Primer sets for synthesising polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 652 AA;

Query Match 97.7%; Score 3504; DB 4; Length 652;
Best Local Similarity 98.2%; Pred. No. 1.9e-263;
Matches 652; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

Qy 1 MERPEEGKQPPPPQPGRLRLGAECEGEPHVLRLKREWTIGRRGCDLSPFNKLVSGDH 60
Db 1 MERPEEGKQPPPPQPGRLRLGAECEGEPHVLRLKREWTIGRRGCDLSPFNKLVSGDH 60
Qy 61 CRIVVDEKSGQVLTEDTSTGTVINKLKVVKKQTCPLQGTGVILVTRKPEPENVAYLY 120
Db 61 CRIVVDEKSGQVLTEDTSTGTVINKLKVVKKQTCPLQGTGVILVTRKPEPENVAYLY 120
Qy 121 ELSSEKQGMTCQESPEANKENVFHTGKDTSGAGAGRGADPRVPPSSPATQVCFEPQPS 180
Db 121 ELSSEKQGMTCQESPEANKENVFHTGKDTSGAGAGRGADPRVPPSSPATQVCFEPQPS 180
Qy 181 TSDLPPTASASSTEPSAGRRSSSCSGGGGSPKSGSPSVASDEVSSASALPDKRTA 240
Db 169 TSDLPPTASASSTEPSAGRRSSSCSGGGGSPKSGSPSVASDEVSSASALPDKRTA 228
Qy 241 SFSSLEPQDELDLEPVKKKRGDGLDLNGQLLVAQPRNAQTVDHVRAAAGKPKMEE 300
Db 229 SFSSLEPQDELDLEPVKKKRGDGLDLNGQLLVAQPRNAQTVDHVRAAAGKPKMEE 288
Qy 301 TLTCTICQDILLHDCVSLQPCMTFCAACYSGWMSRSLCPTCRCPVERICKNHLNVLVE 360
Db 289 TLTCTICQDILLHDCVSLQPCMTFCAACYSGWMSRSLCPTCRCPVERICKNHLNVLVE 348
Qy 361 AYLIQHPDKSRSEEDVQSDMARNKITQDMLQPKVRSFDESGSDDLLELSDVDSESD 420
Db 349 AYLIQHPDKSRSEEDVQSDMARNKITQDMLQPKVRSFDESGSDDLLELSDVDSESD 408
Qy 421 ISQPVVVCRCPEYRQAAPPHCPAPEGEPGAQALGDAPSTVSLTAVQDYVCPLOQ 480
Db 409 ISQPVVVCRCPEYRQAAPPHCPAPEGEPGAQALGDAPSTVSLTAVQDYVCPLOQ 468
Qy 481 SHALCTCCFQPMDDRAERQDPRAVAPQCAVCLQPFCHLYWGCTRTGCGCLAPFCELN 540
Db 469 SHALCTCCFQPMDDRAERQDPRAVAPQCAVCLQPFCHLYWGCTRTGCGCLAPFCELN 528
Qy 541 LGDKCLDGLVNNNSYEDILKNYLATRGLTWKNNLTESLVALQRGVFLSDYRVGTDL 600
Db 529 LGDKCLDGLVNNNSYEDILKNYLATRGLTWKNNLTESLVALQRGVFLSDYRVGTDL 588
Qy 601 CYCCGLRSFRELTYQYRQNIIPASELPVAVTSRPDCYWGRCRTQVKAHMKFNHICEQT 660
Db 589 CYCCGLRSFRELTYQYRQNIIPASELPVAVTSRPDCYWGRCRTQVKAHMKFNHICEQT 648
Qy 661 RPKN 664
Db 661 RPKN 664

Db 649 RFXN 652

RESULT 4

AB97233

ID ABB97233 standard; protein; 623 AA.

XX ABB97233;

AC ABB97233;

DT 27-JUN-2002 (first entry)

XX Novel human protein SEQ ID NO: 501.

DE Human, antianemic; vulnary; antiinflammatory; immunomodulator;

XX antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;

KW neuroprotective; antiparkinsonian; protein therapy; EST;

XX expressed sequence tag.

OS Homo sapiens.

XX WO200222660-A2.

PN 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US026015.

PF 11-SEP-2000; 2000US-00659671.

PR (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

PI WPI: 2002-292408/33.

DR N-FSDB; ABN32419.

XX An isolated polynucleotide for treating diseases associated with its

PT encoded polypeptide such as cancer and multiple sclerosis.

PS Example 2; SEQ ID NO 501; 509pp; English.

XX The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibin e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke

CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.

CC rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention

XX

SQ Sequence 623 AA;

Query Match 88.4%; Score 3168.5; DB 5; Length 623;

Best Local Similarity 88.7%; Fred. NO. 2.2e-237;

Matches 595; Conservative 5; Mismatches 16; Indels 55; Gaps 2;

QY 1 MERPEGKQSPPPQPPWGRLLRLGAEGEPHVLRLKREWTIGRRGCDLSFPPNKLVSQDH 60

DB 1 MERPEGKQSPPPQPPWGRLLRLGAEGEPHVLRLKREWTIGRRGCDLSFPPNKLVSQDH 60

QY 61 CRIVDEKSGQVLTEDTSGTVINKLVKKQTCPLQTDVLYLVYRKNEPEHNVAYLY 120

DB 61 CRIVDEKSGQVLTEDTSGTVINKLVKKQTCPLQTDVLYLVYRKNEPEHNVAYLY 120

QY 121 ESLSEKQGMTOESFEANKENVFHTKDTSGAGARGADPRVPPSPATQVCFEPQPTS 180

DB 121 ESLSEKQGMTOESFEANKENVFHTKDTSGAGARGADPRVPPSPATQVCFEPQPTS 180

QY 181 TSLDFFPTASA-----SSTPEPAGRRERSSSCGSGGGISPKGSGPSVASDEVSSPASA 233

DB 136 ---MVPCCAQAAGKLGLGSDPTLASQSVITVITGSGGGISPKGSGPSVASDEVSSPASA 192

QY 234 LPDRKTAASFSSLEPQOEDLEPVKKQMRGGDLDLNGQLLVQAPRRNAQTVHEDVRAAG 293

DB 193 LPDRKTAASFSSLEPQOEDLEPVKKQMRGGDLDLNGQLLVQAPRRNAQTVHEDVRAAG 252

QY 294 KPDKMEETLTCTIICQDLLHDCVSLQPCMTFCAACYSGWMSRSSLCPTCRCPVERICKNH 353

DB 253 KPDKMEETLTCTIICQDLLHDCVSLQPCMTFCAACYSGWMSRSSLCPTCRCPVERICKNH 312

QY 354 ILNNLVEAYLIQHPDKSRSEEDVQSDARKKITQDMLQPKVRRSFDEGSSDLELSD 413

DB 313 ILNNLVEAYLIQHPDKSRSEEDVQSDARKKITQDMLQPKVRRSFDEGSSDLELSD 372

QY 414 VDESSDISQPYVVCQCEYRRAQAPPHCPAPEGEPGAPQALGAPSTSVSLTTAVQD 473

DB 373 VDESSDISQPYVVCQCEYRRAQAPPHCPAPEGEPGAPQALGAPSTSVSLTTAVQD 432

QY 474 YVCPLOGSHALCTCCFPMPDRRAEREQDPRVAPQCAVCLQPFCHLYMGCTRTGCGCL 533

DB 433 YVCPLOGSHALCTCCFPMPDRRAEREQDPRVAPQCAVCLQPFCHLYMGCTRTGCGCL 492

QY 534 APFCELNLGDKCLDGVNLNNSYESDILKNYLATRGITWKNMLTESLVALQGVFLLSDYR 593

DB 493 APFCELNLGDKCLDGVNLNNSYESDILKNYLATRGITWKNMLTESLVALQGVFLLSDYR 552

QY 594 VTGDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRPPDCYWGRCNRTQVKAHAMKF 653

DB 553 VTGDTVLCYCCGLRSFRELTYQYRONIPASELPVAVTSRPPDCYWGRCNRTQVKAHAMKF 612

QY 654 NHICEQTRFKN 664

DB 613 NHICEQTRFKN 623

RESULT 5

AB93182

ID AAB93182 standard; protein; 623 AA.

XX AAB93182;

AC AAB93182;

DT 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:12128.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.

KW Homo sapiens.

XX EP1074617-A2.

PN 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 12128; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 623 AA;

Query Match 88.2%; Score 3162.5; DB 4; Length 623;
Best Local Similarity 88.5%; Pred. No. 6.5e-237;
Matches 594; Conservative 5; Mismatches 17; Indels 55; Gaps 2;

QY 1 MERPEGKQSPPPQWGRLLRLGAEGBEPHVLRLKKEWTIGRRGCDLSFPFNKLVSQDH 60
DB 1 MERPEGKQSPPPQWGRLLRLGAEGBEPHVLRLKKEWTIGRRGCDLSFPFNKLVSQDH 60

QY 61 CRIVDEKSGQVLEDTSGTVINKLVKKQTCPLQGDVILVYRNKEPHNVAVLY 120
DB 61 CRIVDEKSGQVLEDTSGTVINKLVKKQTCPLQGDVILVYRNKEPHNVAVLY 120

QY 121 ESLSEKQGMTOESFEANKENVFHTGTDGAGAGRADPRVPSPATQVCFEFPQSPST 180
DB 121 ESLSEKQGMTOESFE ----- 135

QY 181 TSLDFTTASA-----SSTEPSAGRRSSSCGGGGISPKGSPSVASDEVSSFASA 233
DB 136 ---MVPCVAAQAGLKLLGSSDPTLASQSIIVTGGGGISPKGSPSVASDEVSSFASA 192

QY 234 LPDRKTASFSSLPQDLEPVKKXRGDGLDLNGQLLVQAQPRNAQTVHEDVRAAAG 293
DB 193 LPDRKTASFSSLPQDLEPEKXMRGDDGLDLNGQLLVQAQPRNAQTVHEDVRAAAG 252

QY 294 KPDKMBETLTCIIQDILLHDCVSLQPMHTFCAACYSQWMSRSLCPTCRCPVERICKNH 353
DB 253 KPDKMBETLTCIIQDILLHDCVSLQPMHTFCAACYSQWMSRSLCPTCRCPVERICKNH 312

QY 354 ILNNLVEAYLIQHPDKSRSEEDVQSMARKITQDMLQPKVRRSFSDEGSSSEDLLELSD 413
DB 313 ILNNLVEAYLIQHPDKSRSEEDVQSMARKITQDMLQPKVRRSFSDEGSSSEDLLELSD 372

QY 414 VDESSDISQPVYVQCQPEYRQAQPPHCPAPEGEPQALGAPSTSVSLTAVQD 473
DB 373 VDESSDISQPVYVQCQPEYRQAQPPHCPAPEGEPQALGAPSTSVSLTAVQD 432

QY 474 YVCPLOQSHALTCCTCPQMPDRAERQDPRVAPQCAVCLQPFCHLYMGCTRTGCGCL 533
DB 433 YVCPLOQSHALTCCTCPQMPDRAERQDPRVAPQCAVCLQPFCHLYMGCTRTGCGCL 492

QY 534 APFCELNLGDKLDGVLNNNSYESDILKNYLATRGLTWKMLTESLVALQGVFLLSYDR 593
DB 493 APFCELNLGDKLDGVLNNNSYESDILKNYLATRGLTWKMLTESLVALQGVFLLSYDR 552

QY 594 VTGDTVLICYCGLRSEFLTYQYRONIPASELPAVTSRPPDCYGNCRCTQVKAHMAKF 653
DB 553 VTGDTVLICYCGLRSEFLTYQYRONIPASELPAVTSRPPDCYGNCRCTQVKAHMAKF 612

QY 654 NHICEQTRFKN 664
DB 613 NHICEQTRFKN 623

RESULT 6
AAU15856
ID AAU15856 standard; protein; 426 AA.
XX AAU15856;
AC AAU15856;
XX
DT 07-NOV-2001 (first entry)
DE Human novel secreted protein, Seq ID 809.
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiac; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198113P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226888P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 14-SEP-2000; 2000US-0233223P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-023484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 03-DEC-2000; 2000US-0251030P.
PR 03-DEC-2000; 2000US-0251988P.
PR 03-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 03-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WFI; 2001-488783/53.
XX N-PSDB; AAS25843.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 809; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Query Match 54.5%; Score 1954.5; DB 4; Length 426;
Best Local Similarity 74.1%; Pred. No. 3.3e-143;
Matches 389; Conservative 5; Mismatches 22; Indels 109; Gaps 5;
OY 1 MERPECKQSPFPQPGWGLRLGAEAGEPHVLRKREWTIGRRGCDLSPFNKLVSGDH 60
DB 3 MERPECKQSPFPQPGWGLRLGAEAGEPHVLRKREWTIGRRGCDLSPFNKLVSGDH 62
OY 61 CRIVDEKSGVLEDTSTSTGTVINKLVKVKCTCPLGTGDIYLYVRKNEPHNVAYL 120

Db 63 CRIWDEKSGQVLTEDTSTGTVINKLVKKVQKTCPLQGTGDIYLVYRKNPEHR----- 117
QY 121 ELSLSEKQGTQTSFENKENVPHGTQKTSAGAGRGADPRVPSPSPATQVCFEFPQPTS 180
Db 118 ----- 117
QY 181 TSDLFPTASASTPSPACGRSSSCGSGGGISPKGSGFVSASDEVSSFASALPDKRTA 240
Db 118 -----SGGGISPKGSGFVSASDEVSSFASALPDKRTA 150
QY 241 SPSLEPQOEDLEPVKKQKRGDGLDINGOLLVQAPRPAQTVHEDVRAAGKPKMEE 300
Db 151 SPSLEPQOEDLEPVKKQKRGDGLDINGOLLVQAPRPAQTVHEDVRAAGKPKMEE 210
QY 301 TLTCICQDLDHDCVSLQPCMTFFCAACYSGWNRSSLCPTCRCPVERICKKHILNLA 360
Db 211 TLTCICQDLDHDCVSLQPCMTFFCAACYSGWNRSSLCPTCRCPVERICKKHILNLA 270
QY 361 AYLIQHPDKSRSEEDVQMDARKKITQDMLQPKVRSPFDEGSSDLELSDVDSSESD 420
Db 271 AYLIQHPDKSRSEEDVQMDARKKITQDMLQPKVRSPFDEGSSDLELSDVDSSESD 330
QY 421 ISQPVVCRQCPRYRQAQPPHCPAPEGEPGAQALGAPSTSVSLTTAVQDYVCPLQG 480
Db 331 ISQPVVCRQCPRYRQAQPPHCPAPEGEPGAQALGAPSTSVSLTT-----VRI 382
QY 481 SHALC---TCCFQMPDRAREQDPRVAPQ-----QCAVCLQPF 517
Db 383 TCALCKEATPC-APAASAHAPESGTRTGPRVXXLNKCKXGLLQAF 426
RESULT 7
ABU54925
ID ABU54925 standard; protein; 426 AA.
XX
AC ABU54925;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polypeptide #12.
XX
KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
FN US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-00764864.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180282P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 08-SEP-2000; 2000US-0229513P.
PR 21-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 27-SEP-2000; 2000US-0234997P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239325P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 01-NOV-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR N-PSDB; ABX73184.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 11; SEQ ID NO 809; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 426 AA;

Query Match 54.5%; Score 1954.5; DB 6; Length 426;
Best Local Similarity 74.1%; Pred. NO. 3-3e-143;
Matches 389; Conservative 5; Mismatches 22; Indels 109; Gaps 5;

QY	1	MERPEEGKQSPPPQFWGRLRLRGAEEGEPHVLLLRKEWTIGRRRCGLDLSFSPNKLVSGDH	60
Db	3	MERPEEGKQSPPPQFWGRLRLRGAEEGEPHVLLLRKEWTIGRRRCGLDLSFSPNKLVSGDH	62
QY	61	CRIVVDEKSGGVTTLEDTSGTGWINKLVKKVKTCTPLQTGDVYLIVYRKNEPEHNVALY	120
Db	63	CRIVVDEKSGGVTTLEDTSGTGWINKLVKKVKTCTPLQTGDVYLIVYRKNEPEHR	117
QY	121	ESLSEKQMTQESFEANKENVFHGKTDSGAGAGRGADPRVPSPATQVCPEEPQPS	180
Db	118	-----	117
QY	181	TSDLPTPTASASTPSPAGRERSSCGSGGGISPKGSGPSVASDEVSSFASALPDKTA	240
Db	118	-----SGGGISPKGSGPSVASDEVSSFASALPDKTA	150
QY	241	SFSSLEPQDQDLDFPVKKMGDGLDLNGQLLVAQPRRNAQTVHEDVAAAGKPKMEE	300
Db	151	SFSSLEPQDQDLDFPVKKMGDGLDLNGQLLVAQPRRNAQTVHEDVAAAGKPKMEE	210
QY	301	TLTCTIICODLLHDCVSLQPCMHTECAACYSGWMERSSLCPTCRCPVERICKHILNNLVE	360
Db	211	TLTCTIICODLLHDCVSLQPCMHTECAACYSGWMERSSLCPTCRCPVERICKHILNNLVE	270
QY	361	AYLIIOHPDKSRSEEDVQSMARNKITQDMLQPKVRRSFDEEGSSDLLELSDVDSSESD	420
Db	271	AYLIIOHPDKSRSEEDVQSMARNKITQDMLQPKVRRSFDEEGSSDLLELSDVDSSESD	330
QY	421	ISQVYVCKRCPEYRRAAQPHCPAPGEGEPGALGDAPSIVSLTTAVQDYVCPKQG	480
Db	331	ISQVYVCKRCPEYRRAAQPHCPAPGEGEPGALGDAPPTSVSLTT-----VRI	382
QY	481	SHALC---TCCFQPMFDRRAAREQDPRVAPQ-----QCAVCLQPF	517
Db	383	TCALCKEATPC-APAASAHAPESGTRTGPVXXLNKCXGLLQAF	426

RESULT 8
AA008972
ID AAC08972 standard; protein; 269 AA.

RESULT 8	
AAO08972	
ID	AAO08972 standard; protein; 269 AA.
XX	
XX	AAO08972;
XX	
XX	AC
XX	
XX	06-NOV-2001 (first entry)
DT	
XX	Human polypeptide SEQ ID NO 22864.
XX	
XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;
KW	nervous system disorders; arthritis; inflammation.
XX	
XX	
OS	Homo sapiens.
XX	
PN	WO200164835-A2.
XX	
XX	07-SEP-2001.
FD	
XX	
XX	26-FEB-2001; 2001WO-US004927.
XX	
XX	28-FEB-2000; 2000US-00515126.
PR	
PR	18-MAY-2000; 2000US-00577409.
XX	
XX	(HYSE-) HYSEQ INC.
PA	
XX	
XX	Tang YT, Liu C, Drmanac RT;
PI	
XX	
XX	WPI; 2001-514838/56.
DR	
DR	N-PSDB; AAI88903.
XX	
XX	Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT	and treating e.g. leukemia, inflammation and immune disorders.
PT	
XX	

Claim 20; SEQ ID NO 22864; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AAI79941-AAI93841) and encoded proteins (AAO0010-AAO13910) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 269 AA;
SQ

AA		Sequence 269 AA;			
SQ					
		Query Match	38.0%; Score 1364; DB 4; Length 269;		
		Best Local Similarity	99.6%; Pred. NO. 1.5e-97;		
		Matches 242; Conservative	1; Mismatches 0; Indels 0; Gaps 0;		
QY	422	SQPVVCRQCPEYRRQAAPPHCPAGEGPGAPALGDAPSVSLSLTTAVODIVCPLQGS	481		
DB	27	SQPVVCRQCPEYRRAAOPHPCPAGEGPGAPALGDAPSVSLSLTTAVODIVCPLQGS	86		
QY	482	HALLCTCCFQMPDPRAREERODPVAPOOCAVCLQQPFCHLYWGCTRITGCYGCLAFFCELNL	541		
DB	87	HALLCTCCFQMPDPRAREERODPVAPOOCAVCLQQPFCHLYWGCTRITGCYGCLAFFCELNL	146		
QY	542	GDKLDGLVNNNYESDLKNIYLATRLGWKNLTESLVALRGVFLLSDRYVTGDTVLIC	601		
DB	147	GDKLDGLVNNNYESDLKNIYLATRLGWKNLTESLMALRGVFLLSDRYVTGDTVLIC	206		
QY	602	YCCGLRSFREITYQRNIPASELPVAVTSRPDCYWGRNCRTQVKAHAMKENHCETR	661		
DB	207	YCCGLRSFREITYQRNIPASELPVAVTSRPDCYWGRNCRTQVKAHAMKENHCETR	266		
QY	662	PKN 664 			
DB	267	PKN 269 			
		RESULT 9			
ID	AUUI6317	standard: protein; 230 AA.			

RESULT 9	
AAU16317	AAU16317
ID	AAU16317 standard; protein; 230 AA.
XX	
XX	AAU16317;
AC	
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human novel secreted protein, Seq ID 1270.
XX	
XX	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW	cardiant; vasotropic; cerbroprotective; neurotropic; neuroprotective;
KW	antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW	cerebral ischaemia; angiogenesis; nervous system disorder;
KW	Alzheimer's disease; infection; ocular disorder; corneal infection;
KW	wound healing; epithelial cell proliferation; skin ageing; food additive;
KW	preservative; antiproliferative.
OS	Homo sapiens.
XX	
PN	WO20015322-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US0001341.
XX	
XX	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	

24-FEB-2000; 2000US-0184664P. PR
02-MAR-2000; 2000US-0186350P. PR
16-MAR-2000; 2000US-0189874P. PR
17-MAR-2000; 2000US-0190076P. PR
18-APR-2000; 2000US-0198123P. PR
19-MAY-2000; 2000US-0205515P. PR
07-JUN-2000; 2000US-0209467P. PR
28-JUN-2000; 2000US-0214886P. PR
30-JUN-2000; 2000US-0215135P. PR
07-JUL-2000; 2000US-0216647P. PR
07-JUL-2000; 2000US-0216880P. PR
11-JUL-2000; 2000US-0217487P. PR
11-JUL-2000; 2000US-0217436P. PR
14-JUL-2000; 2000US-0218290P. PR
26-JUL-2000; 2000US-0220963P. PR
26-JUL-2000; 2000US-0220964P. PR
14-AUG-2000; 2000US-0224518P. PR
14-AUG-2000; 2000US-0224519P. PR
14-AUG-2000; 2000US-0225213P. PR
14-AUG-2000; 2000US-0225214P. PR
14-AUG-2000; 2000US-0225266P. PR
14-AUG-2000; 2000US-0225267P. PR
14-AUG-2000; 2000US-0225268P. PR
14-AUG-2000; 2000US-0225270P. PR
14-AUG-2000; 2000US-0225447P. PR
14-AUG-2000; 2000US-0225757P. PR
14-AUG-2000; 2000US-0225758P. PR
14-AUG-2000; 2000US-0225759P. PR
18-AUG-2000; 2000US-0226279P. PR
22-AUG-2000; 2000US-0226661P. PR
22-AUG-2000; 2000US-0226668P. PR
22-AUG-2000; 2000US-0227182P. PR
22-AUG-2000; 2000US-0227009P. PR
30-AUG-2000; 2000US-0228924P. PR
01-SEP-2000; 2000US-0229287P. PR
01-SEP-2000; 2000US-0229343P. PR
01-SEP-2000; 2000US-0229344P. PR
01-SEP-2000; 2000US-0229345P. PR
05-SEP-2000; 2000US-0229509P. PR
05-SEP-2000; 2000US-0229513P. PR
06-SEP-2000; 2000US-0230437P. PR
06-SEP-2000; 2000US-0230438P. PR
08-SEP-2000; 2000US-0231242P. PR
08-SEP-2000; 2000US-0231243P. PR
08-SEP-2000; 2000US-0231244P. PR
08-SEP-2000; 2000US-0231413P. PR
08-SEP-2000; 2000US-0231414P. PR
08-SEP-2000; 2000US-0232080P. PR
08-SEP-2000; 2000US-0232081P. PR
12-SEP-2000; 2000US-0231968P. PR
14-SEP-2000; 2000US-0232397P. PR
14-SEP-2000; 2000US-0232398P. PR
14-SEP-2000; 2000US-0232399P. PR
14-SEP-2000; 2000US-0232400P. PR
14-SEP-2000; 2000US-0232401P. PR
14-SEP-2000; 2000US-0233063P. PR
14-SEP-2000; 2000US-0233064P. PR
14-SEP-2000; 2000US-0233065P. PR
21-SEP-2000; 2000US-0234223P. PR
21-SEP-2000; 2000US-0234274P. PR
25-SEP-2000; 2000US-0234997P. PR
25-SEP-2000; 2000US-0234998P. PR
26-SEP-2000; 2000US-0235484P. PR
27-SEP-2000; 2000US-0235834P. PR
27-SEP-2000; 2000US-0235836P. PR
29-SEP-2000; 2000US-0236327P. PR
29-SEP-2000; 2000US-0236367P. PR
29-SEP-2000; 2000US-0236368P. PR
29-SEP-2000; 2000US-0236369P. PR
29-SEP-2000; 2000US-0236370P. PR
02-OCT-2000; 2000US-0236802P. PR
02-OCT-2000; 2000US-0237037P. PR
02-OCT-2000; 2000US-0237038P. PR
02-OCT-2000; 2000US-0237039P. PR
02-OCT-2000; 2000US-0237040P. PR
13-OCT-2000; 2000US-0239935P. PR
13-OCT-2000; 2000US-0239937P. PR
20-OCT-2000; 2000US-0240960P. PR
20-OCT-2000; 2000US-0241221P. PR
20-OCT-2000; 2000US-0241785P. PR
20-OCT-2000; 2000US-0241786P. PR
20-OCT-2000; 2000US-0241808P. PR
20-OCT-2000; 2000US-0241809P. PR
20-OCT-2000; 2000US-0241826P. PR
01-NOV-2000; 2000US-0244617P. PR
08-NOV-2000; 2000US-0246474P. PR
08-NOV-2000; 2000US-0246475P. PR
08-NOV-2000; 2000US-0246478P. PR
08-NOV-2000; 2000US-0246477P. PR
08-NOV-2000; 2000US-0246523P. PR
08-NOV-2000; 2000US-0246524P. PR
08-NOV-2000; 2000US-0246525P. PR
08-NOV-2000; 2000US-0246526P. PR
08-NOV-2000; 2000US-0246527P. PR
08-NOV-2000; 2000US-0246528P. PR
08-NOV-2000; 2000US-0246532P. PR
08-NOV-2000; 2000US-0246609P. PR
08-NOV-2000; 2000US-0246610P. PR
08-NOV-2000; 2000US-0246611P. PR
08-NOV-2000; 2000US-0246613P. PR
17-NOV-2000; 2000US-0249207P. PR
17-NOV-2000; 2000US-0249208P. PR
17-NOV-2000; 2000US-0249209P. PR
17-NOV-2000; 2000US-0249210P. PR
17-NOV-2000; 2000US-0249211P. PR
17-NOV-2000; 2000US-0249212P. PR
17-NOV-2000; 2000US-0249244P. PR
17-NOV-2000; 2000US-0249245P. PR
17-NOV-2000; 2000US-0249245P. PR
17-NOV-2000; 2000US-0249264P. PR
17-NOV-2000; 2000US-0249265P. PR
17-NOV-2000; 2000US-0249297P. PR
17-NOV-2000; 2000US-0249299P. PR
17-NOV-2000; 2000US-0249300P. PR
01-DEC-2000; 2000US-0250160P. PR
01-DEC-2000; 2000US-0250391P. PR
05-DEC-2000; 2000US-0251030P. PR
05-DEC-2000; 2000US-0251988P. PR
05-DEC-2000; 2000US-0256719P. PR
06-DEC-2000; 2000US-0251479P. PR
08-DEC-2000; 2000US-0251868P. PR
08-DEC-2000; 2000US-0251869P. PR
08-DEC-2000; 2000US-0251989P. PR
08-DEC-2000; 2000US-0251990P. PR
11-DEC-2000; 2000US-0254097P. PR
05-JAN-2001; 2001US-0259678P. PR
(HUMA-) HUMAN GENOME SCI INC. PA
Rosen CA, Barash SC, Ruben SM; XX
WFI; 2001-488783/53. XX
N-PSDB; AAS26304. XX
New nucleic acid molecules encoding 461 human secreted proteins for PT
diagnosing, preventing, treating or ameliorating medical conditions and PT
used as food additives or preservatives. XX

PS Claim 11; SEQ ID NO 1270; 980pp; English.

XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 23.5%; Score 843.5; DB 4; Length 230;
Best Local Similarity 76.3%; Pred. No. 3.5e-57;
Matches 167; Conservative 2; Mismatches 19; Indels 31; Gaps 4;

QY 319 PCMHTECAACYSQWMSRLCTPCPVERICKNHLNNLVEAYLIQHPDKSRSEEDVQS 378
Db 5 PCMHTECAACYSQWMSRLCTPCPVERICKNHLNNLVEAYLIQHPDKSRSEEDVQS 64

QY 379 MDARNKITQMLQPKVRRSFSEDEGSDELLESDVDESSDISQPVVVCRCPPYRQA 438
Db 65 MDARNKITQMLQPKVRRSFSEDEGSDELLESDVDESSDISQPVVVCRCPPYRQA 124

QY 439 AQPHPCPAPEGEPGAPQALGDAPSTSVLTAVQDVCPLOGSHALCTCCFQMP- 493
Db 125 AQPHPCPAPEGEPGAPQALGDAPSTSVLTAVQDVCPLOGSHALCTCCFQMP- 173

QY 494 --DRRAE--REOPRVA-----PQCAVCLQPF 517
Db 174 SAHARPESGREQDPRXALXXPPVLGLHPDLRLILAPF 212

RESULT 10
ABUS5386
ID ABUS5386 standard; protein; 230 AA.
XX ABUS5386;
XX AC ABUS5386;
XX DE 18-MAR-2003 (first entry)
XX Human novel polypeptide #473.
XX Human; neural disorder; immune system disorder; renal disorder;
XX muscular disorder; respiratory disease; reproductive disorder;
XX gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX hyperproliferative disorder; inflammatory disease; allergic reaction;
XX blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
XX haemostatic; antiarteriosclerotic.
XX Homo sapiens.
XX OS
XX US2002132753-A1.
XX PD 19-SEP-2002.
XX PF 17-JAN-2001; 2001US-00764864.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-021680P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 08-SEP-2000; 2000US-0229513P.
PR 21-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234232P.
PR 25-SEP-2000; 2000US-0234274P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
PI Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-147444/14.
XX N-PSDB; ABX73645.
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX Claim 11; SEQ ID NO 1270; 402pp; English.
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.

CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 230 AA;

Query Match 23.5%; Score 843.5; DB 6; Length 230;
Best Local Similarity 76.3%; Pred. No. 3.5e-57;
Matches 167; Conservative 2; Mismatches 19; Indels 31; Gaps 4;

Qy 319 PCWHTFCAACYSYGMWSSSLCTCRCPVERICKHILNNLVEAYLIQHDPKSRSEEDVQS 378
Db 5 PCWHTFCAACYSYGMWSSSLCTCRCPVERICKHILNNLVEAYLIQHDPKSRSEEDVQS 64

Qy 379 MDARKKITQDMLQPKVRSFSDSEGSSEDLLELSDVDSESDISQPVVVCRCQPEYRROA 438
Db 65 MDARKKITQDMLQPKVRSFSDSEGSSEDLLELSDVDSESDISQPVVVCRCQPEYRROA 124

Qy 439 AQPPECAPGEGPGAPALGDAPSTSVSLTAVQDYVCPQGGSHALCTCCFQMP-----493
Db 125 AQPPECAPGEGPGAPALGDAPSTSVSLTAV-----RITCALCKEATPCAPAA 173

Qy 494 --DRAE--REQPRVA-----POCCAVCIQPF 517
Db 174 SAHARPSGREGQPRXALXXXPPVLGLHPDLRLRLAPF 212

RESULT 11
ID ADA54231 standard; protein; 128 AA.
XX
AC ADA54231;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 1799.
XX
KW Cystostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
FN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
XX
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
FI Isogai T, Sugiyama T, Otsuki T, Wakanatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA52592.
XX
FT New polynucleotides encoding full-length polypeptides, e.g. secretory
FT and/or membrane proteins, useful for developing medicines for diseases in
FT which the gene is involved, or as target molecules for gene therapy.
XX

PS Claim 14; SEQ ID NO 1799; 205pp; English.
XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 128 AA;

Query Match 15.8%; Score 566; DB 6; Length 128;
Best Local Similarity 99.0%; Pred. No. 5.9e-36;
Matches 103; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 561 KNYLATRGLTWKMLTESLVALQSGVFLLSYRTGDTVLCYCCGLRSFRLTYQYRONI 620
Db 25 QNYLATRGLTWKMLTESLVALQSGVFLLSYRTGDTVLCYCCGLRSFRLTYQYRONI 84

Qy 621 FASELPVAVTSRPPDCYWGRCNCRTOVKAHAMKFNHICEQTRFKN 664
Db 85 FASELPVAVTSRPPDCYWGRCNCRTOVKAHAMKFNHICEQTRFKN 128

RESULT 12
ID AAU16156 standard; protein; 92 AA.
XX
AC AAU16156;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1109.
XX
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
FN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190078P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 11-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.

PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225269P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0225273P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0225681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0225686P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-SEP-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249214P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249216P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0231142P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231143P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231144P.	PR	17-NOV-2000;	2000US-0249246P.
PR	08-SEP-2000;	2000US-0231145P.	PR	17-NOV-2000;	2000US-0249254P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249255P.
PR	08-SEP-2000;	2000US-0232081P.	PR	17-NOV-2000;	2000US-0249256P.
PR	12-SEP-2000;	2000US-0231368P.	PR	17-NOV-2000;	2000US-0249257P.
PR	14-SEP-2000;	2000US-0232397P.	PR	17-NOV-2000;	2000US-0249259P.
PR	14-SEP-2000;	2000US-0232398P.	PR	17-NOV-2000;	2000US-0249259P.
PR	14-SEP-2000;	2000US-0232399P.	PR	17-NOV-2000;	2000US-0249259P.
PR	14-SEP-2000;	2000US-0232400P.	PR	17-NOV-2000;	2000US-0249259P.
PR	14-SEP-2000;	2000US-0233063P.	PR	17-NOV-2000;	2000US-0249259P.
PR	14-SEP-2000;	2000US-0233064P.	PR	17-NOV-2000;	2000US-0249259P.
PR	14-SEP-2000;	2000US-0233065P.	PR	17-NOV-2000;	2000US-0249259P.
PR	21-SEP-2000;	2000US-0234223P.	PR	17-NOV-2000;	2000US-0249259P.
PR	21-SEP-2000;	2000US-0234274P.	PR	17-NOV-2000;	2000US-0249259P.
PR	25-SEP-2000;	2000US-0234957P.	PR	17-NOV-2000;	2000US-0249259P.
PR	25-SEP-2000;	2000US-0234998P.	PR	17-NOV-2000;	2000US-0249259P.
PR	25-SEP-2000;	2000US-0235484P.	PR	17-NOV-2000;	2000US-0249259P.
PR	27-SEP-2000;	2000US-0235834P.	PR	17-NOV-2000;	2000US-0249259P.
PR	27-SEP-2000;	2000US-0235836P.	PR	17-NOV-2000;	2000US-0249259P.
PR	29-SEP-2000;	2000US-0236327P.	PR	17-NOV-2000;	2000US-0249259P.
PR	29-SEP-2000;	2000US-0236367P.	PR	17-NOV-2000;	2000US-0249259P.
PR	29-SEP-2000;	2000US-0236368P.	PR	17-NOV-2000;	2000US-0249259P.
PR	29-SEP-2000;	2000US-0236370P.	PR	17-NOV-2000;	2000US-0249259P.
PR	29-SEP-2000;	2000US-0236370P.	PR	17-NOV-2000;	2000US-0249259P.
PR	02-OCT-2000;	2000US-0236380P.	PR	17-NOV-2000;	2000US-0249259P.
PR	02-OCT-2000;	2000US-0237037P.	PR	17-NOV-2000;	2000US-0249259P.
PR	02-OCT-2000;	2000US-0237038P.	PR	17-NOV-2000;	2000US-0249259P.
PR	02-OCT-2000;	2000US-0237039P.	PR	17-NOV-2000;	2000US-0249259P.
PR	02-OCT-2000;	2000US-0237040P.	PR	17-NOV-2000;	2000US-0249259P.
PR	13-OCT-2000;	2000US-0239335P.	PR	17-NOV-2000;	2000US-0249259P.
PR	13-OCT-2000;	2000US-0239337P.	PR	17-NOV-2000;	2000US-0249259P.
PR	20-OCT-2000;	2000US-0240960P.	PR	17-NOV-2000;	2000US-0249259P.
PR	20-OCT-2000;	2000US-0241221P.	PR	17-NOV-2000;	2000US-0249259P.
PR	20-OCT-2000;	2000US-0241785P.	PR	17-NOV-2000;	2000US-0249259P.
PR	20-OCT-2000;	2000US-0241786P.	PR	17-NOV-2000;	2000US-0249259P.
PR	20-OCT-2000;	2000US-0241787P.	PR	17-NOV-2000;	2000US-0249259P.
PR	20-OCT-2000;	2000US-0241808P.	PR	17-NOV-2000;	2000US-0249259P.
PR	20-OCT-2000;	2000US-0241809P.	PR	17-NOV-2000;	2000US-0249259P.
PR	20-OCT-2000;	2000US-0241826P.	PR	17-NOV-2000;	2000US-0249259P.
PR	01-NOV-2000;	2000US-0244647P.	PR	17-NOV-2000;	2000US-0249259P.
PR	08-NOV-2000;	2000US-0246475P.	PR	17-NOV-2000;	2000US-0249259P.
PR	08-NOV-2000;	2000US-0246476P.	PR	17-NOV-2000;	2000US-0249259P.
PR	08-NOV-2000;	2000US-0246477P.	PR	17-NOV-2000;	2000US-0249259P.
PR	08-NOV-2000;	2000US-0246478P.	PR	17-NOV-2000;	2000US-0249259P.
PR	08-NOV-2000;	2000US-0246523P.	PR	17-NOV-2000;	2000US-0249259P.
PR	08-NOV-2000;	2000US-0246524P.	PR	17-NOV-2000;	2000US-0249259P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.

N-P8DB; AAS26143.

New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 11; SEQ ID NO 1109; 980pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to

CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 13.2%; Score 474; DB 4; Length 92;
Best Local Similarity 98.9%; Pred. No. 5.3e-29;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MERPEEGKQSPFPQPGWGLRLGAEAGEPHVLLRKREWTIGRRGCDLSPFNKLVSGDH 60
DB 3 MERPEEGKQSPFPQPGWGLRLGAEAGEPHVLLRKREWTIGRRGCDLSPFNKLVSGDH 62
QY 61 CRIVVDEKSGQVLTEDTSTSGTVINKLVK 90
DB 63 CRIVVDEKSGQVLTEDTSTSGTVINKLVK 92

RESULT 13
ABUS5225
ID ABUS5225 standard; protein; 92 AA.

XX AC ABUS5225;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polypeptide #312.
XX KW Human; neural disorder; immune system disorder; renal disorder;
XX KW muscular disorder; respiratory disease; reproductive disorder;
XX KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX KW hyperproliferative disorder; inflammatory disease; allergic reaction;
XX KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
XX KW haemostatic; antiarteriosclerotic.

XX CS Homo sapiens.
XX PN US2002132753-A1.
XX PD 19-SEP-2002.

XX PF 17-JAN-2001; 2001US-00764864.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.
XX PR 14-AUG-2000; 2000US-0225270P.
XX PR 14-AUG-2000; 2000US-0225447P.
XX PR 14-AUG-2000; 2000US-0225757P.
XX PR 14-AUG-2000; 2000US-0225758P.
XX PR 22-AUG-2000; 2000US-0226688P.
XX PR 30-AUG-2000; 2000US-0228924P.
XX PR 01-SEP-2000; 2000US-0229287P.
XX PR 01-SEP-2000; 2000US-0229343P.
XX PR 01-SEP-2000; 2000US-0229344P.
XX PR 01-SEP-2000; 2000US-0229345P.
XX PR 05-SEP-2000; 2000US-0229509P.
XX PR 05-SEP-2000; 2000US-0229513P.
XX PR 08-SEP-2000; 2000US-0231413P.

PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 12-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX PI Rosen CA, Ruben SM, Barash SC;
XX DR WPI: 2003-147444/14.
XX DR N-PSDB; ABX73484.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,
XX inhibiting or preventing e.g. neural, immune system, muscular,
XX respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
XX renal disorders.

XX Claim 11; SEQ ID NO 1109; 402pp; English.

XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,
XX (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
XX leukaemia), inflammatory reactions (e.g. septic shock, bursitis and
XX appendicitis), allergic reactions and conditions (e.g. asthma), blood
XX related disorders (e.g. thrombosis, atherosclerosis and myocardial
XX infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and
XX ABUS5748 represent human novel polypeptides of the invention

XX Sequence 92 AA;

Query Match 13.2%; Score 474; DB 6; Length 92;
Best Local Similarity 98.9%; Pred. No. 5.3e-29;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MERPEEGKQSPFPQPGWGLRLGAEAGEPHVLLRKREWTIGRRGCDLSPFNKLVSGDH 60
DB 3 MERPEEGKQSPFPQPGWGLRLGAEAGEPHVLLRKREWTIGRRGCDLSPFNKLVSGDH 62
QY 61 CRIVVDEKSGQVLTEDTSTSGTVINKLVK 90
DB 63 CRIVVDEKSGQVLTEDTSTSGTVINKLVK 92

RESULT 14
AAU16157

ID	AAU16157 standard; protein; 99 AA.	PR	08-SEP-2000; 2000US-0231413P
XX	AAU16157;	PR	08-SEP-2000; 2000US-0231414P
XX		PR	08-SEP-2000; 2000US-0232080P
XX		PR	08-SEP-2000; 2000US-0232081P
XX		PR	12-SEP-2000; 2000US-0231968P
XX	07-NOV-2001 (first entry)	PR	14-SEP-2000; 2000US-0232397P
XX	Human novel secreted protein, Seq ID 1110.	PR	14-SEP-2000; 2000US-0232398P
XX		PR	14-SEP-2000; 2000US-0232399P
XX	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;	PR	14-SEP-2000; 2000US-0232400P
KW	cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;	PR	14-SEP-2000; 2000US-0232401P
KW	antibacterial; virucide; fungicide; ophthalmological; vulnerary;	PR	14-SEP-2000; 2000US-0233063P
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;	PR	14-SEP-2000; 2000US-0233064P
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;	PR	14-SEP-2000; 2000US-0233065P
KW	cerebral ischaemia; angiogenesis; nervous system disorder;	PR	21-SEP-2000; 2000US-0234221P
KW	Alzheimer's disease; infection; ocular disorder; corneal infection;	PR	21-SEP-2000; 2000US-0234227P
KW	wound healing; epithelial cell proliferation; skin ageing; food additive;	PR	25-SEP-2000; 2000US-0234998P
KW	preservative; antiproliferative.	PR	26-SEP-2000; 2000US-0235484P
OS	Homo sapiens.	PR	27-SEP-2000; 2000US-0235834P
XX		PR	27-SEP-2000; 2000US-0235836P
XX		PR	29-SEP-2000; 2000US-0236327P
XX		PR	29-SEP-2000; 2000US-0236367P
PN		PR	29-SEP-2000; 2000US-0236368P
XX		PR	29-SEP-2000; 2000US-0236369P
XX		PR	29-SEP-2000; 2000US-0236370P
PD	02-AUG-2001.	PR	02-OCT-2000; 2000US-0236802P
XX	17-JAN-2001; 2001WO-US001341.	PR	02-OCT-2000; 2000US-0237037P
XX		PR	02-OCT-2000; 2000US-0237038P
XX		PR	02-OCT-2000; 2000US-0237039P
PR	31-JAN-2000; 2000US-0179065P.	PR	02-OCT-2000; 2000US-0237040P
PR	04-FEB-2000; 2000US-0180628P.	PR	13-OCT-2000; 2000US-0239935P
PR	24-FEB-2000; 2000US-0184664P.	PR	13-OCT-2000; 2000US-0239937P
PR	02-MAR-2000; 2000US-0186350P.	PR	20-OCT-2000; 2000US-0240960P
PR	16-MAR-2000; 2000US-0189874P.	PR	20-OCT-2000; 2000US-0241221P
PR	17-MAR-2000; 2000US-0190076P.	PR	20-OCT-2000; 2000US-0241785P
PR	18-APR-2000; 2000US-0198123P.	PR	20-OCT-2000; 2000US-0241788P
PR	19-MAY-2000; 2000US-0205151P.	PR	20-OCT-2000; 2000US-0241808P
PR	07-JUN-2000; 2000US-0209467P.	PR	20-OCT-2000; 2000US-0241809P
PR	28-JUN-2000; 2000US-0214986P.	PR	20-OCT-2000; 2000US-0241826P
PR	30-JUN-2000; 2000US-0215135P.	PR	01-NOV-2000; 2000US-0244617P
PR	07-JUL-2000; 2000US-0216647P.	PR	08-NOV-2000; 2000US-0246474P
PR	07-JUL-2000; 2000US-0216980P.	PR	08-NOV-2000; 2000US-0246475P
PR	11-JUL-2000; 2000US-0217487P.	PR	08-NOV-2000; 2000US-0246478P
PR	11-JUL-2000; 2000US-0217496P.	PR	08-NOV-2000; 2000US-0246479P
PR	14-JUL-2000; 2000US-0218290P.	PR	08-NOV-2000; 2000US-0246479P
PR	26-JUL-2000; 2000US-0220963P.	PR	08-NOV-2000; 2000US-0246523P
PR	26-JUL-2000; 2000US-0220964P.	PR	08-NOV-2000; 2000US-0246523P
PR	14-AUG-2000; 2000US-0224518P.	PR	08-NOV-2000; 2000US-0246525P
PR	14-AUG-2000; 2000US-0224519P.	PR	08-NOV-2000; 2000US-0246525P
PR	14-AUG-2000; 2000US-0225213P.	PR	08-NOV-2000; 2000US-0246528P
PR	14-AUG-2000; 2000US-0225214P.	PR	08-NOV-2000; 2000US-0246532P
PR	14-AUG-2000; 2000US-0225266P.	PR	08-NOV-2000; 2000US-0246609P
PR	14-AUG-2000; 2000US-0225267P.	PR	08-NOV-2000; 2000US-0246610P
PR	14-AUG-2000; 2000US-0225268P.	PR	08-NOV-2000; 2000US-0246611P
PR	14-AUG-2000; 2000US-0225270P.	PR	08-NOV-2000; 2000US-0246613P
PR	14-AUG-2000; 2000US-0225447P.	PR	17-NOV-2000; 2000US-0249207P
PR	14-AUG-2000; 2000US-0225475P.	PR	17-NOV-2000; 2000US-0249208P
PR	14-AUG-2000; 2000US-0225758P.	PR	17-NOV-2000; 2000US-0249209P
PR	14-AUG-2000; 2000US-0225759P.	PR	17-NOV-2000; 2000US-0249210P
PR	18-AUG-2000; 2000US-0226279P.	PR	17

PR 17-NOV-2000; 2000US-02492972.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254087P.
PR 03-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX N-PSDB; AAS26144.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1110; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angioneurosis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Query Match 12.1%; Score 432; DB 4; Length 99;
Best Local Similarity 92.5%; Pred. No. 1.1e-25;
Matches 86; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
QY 42 RRGCDLSPFNKLVSDHCRIVVDEKSGQVLTSTGTVINKLKVKKQTCPLQTGD 101
Db 9 RDRGPD--FPXKLVSDHCRIVVDEKSGQVLTSTGTVINKLKVKKQTCPLQTGD 66
QY 102 VIVLVYRNKPEHNVAYLYESLSKQGMTCESF 134
Db 67 VIVLVYRNKPEHNVAYLYESLSKQGMTCESF 99
RESULT 15
AAU16576
ID AAU16576 standard; protein; 99 AA.
XX
XX AAU16576;
AC

XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1529.
XX
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiac; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angioneurosis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
XX WO200155322-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.

PR PR PR PR PR PR PR PR PR PR PR PR PR PR XX PA XX PI XX DR DR XX PT PT PT XX PS XX CC

Rosen CA, Barash SC, Ruben SM;
WPI; 2001-488783/53.
N-PSDB: AAS26563.

Claim 11: SEO ID NO 1529: 980pp: English.

Query Match 12.1%; Score 432; DB 4; Length 99;
Best Local Similarity 92.5%; Pred. No. 1.1e-25;
Matches 86: Conservative 1; Mismatches 4; Indels

42 RRGCDLSPSPNKLVSGBDHCRIVVDEKSGQVLTETSTSGNTVINLKVVKKQTCPLOTGD 101
:
9 RDKGPD--PPXNKLVSGDHCRIWDEKSGQVLTETSTSGNTVINLKVVKKQTCPLOTGD 66

102 VILVVRKNPEHNVAIYLYESISEKGMQTQESF 134

67 VILVVRKNPEHNVAIYLYESISEKGMQTQESF 99

Search completed: May 7, 2004, 14:45:44
Job time : 139.275 secs

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:42:23 ; Search time 44.1732 Seconds
(without alignments)

776.028 Million cell updates/sec

Title: US-10-048-046-3

Perfect score: 3585

Sequence: 1 MERPEEGKQSPPPQPWGRLL.....VKAHAMKFNHICEQTRFKN 664

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /cqn2 6/ptodata/2/iaa/5A COMB.pep:*
```

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pcp.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.per

6: /cgn2_6/ptodata/2/iaa/backfiles1.perf

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	152	4.2	245	2	US-08-897-340-36		Sequence 36, Appl
2	152	4.2	245	3	US-09-252-323-36		Sequence 36, Appl
3	135	3.8	67	1	US-07-945-283-7		Sequence 7, Appl
4	135	3.8	245	2	US-08-897-340-35		Sequence 35, Appl
5	135	3.8	245	3	US-09-252-323-35		Sequence 35, Appl
6	134.5	3.8	826	4	US-09-894-998A-47		Sequence 47, Appl
7	132.5	3.7	2414	1	US-08-227-536-2		Sequence 2, Appl
8	132.5	3.7	2414	5	PC1-US95-04682-2		Sequence 2, Appl
9	131	3.7	1848	3	US-08-296-791-6		Sequence 6, Appl
10	131	3.7	1848	4	US-09-839-999-6		Sequence 6, Appl
11	131	3.7	1848	4	US-10-080-505-6		Sequence 6, Appl
12	131	3.7	1848	5	PC1-US95-10661A-6		Sequence 6, Appl
13	127.5	3.6	989	4	US-08-213-419B-4		Sequence 4, Appl
14	126.5	3.5	933	3	US-08-764-870-14		Sequence 14, Appl
15	126.5	3.5	933	3	US-08-980-115-14		Sequence 14, Appl
16	125.5	3.5	289	4	US-08-213-419B-2		Sequence 2, Appl
17	125	3.5	2482	4	US-09-252-991A-16967		Sequence 16967, A
18	123.5	3.4	1233	4	US-09-688-188B-89		Sequence 89, Appl
19	123.5	3.4	1233	4	US-09-291-417D-89		Sequence 89, Appl
20	123	3.4	395	2	US-08-841-349-9		Sequence 9, Appl
21	123	3.4	395	4	US-09-431-184A-9		Sequence 9, Appl
22	122	3.4	778	6	5198347-4		Patent No. 5198347
23	121.5	3.4	848	1	US-08-045-806-4		Sequence 4, Appl
24	121.5	3.4	848	1	US-08-366-051B-4		Sequence 4, Appl
25	121.5	3.4	1719	2	US-08-459-568-4		Sequence 4, Appl
26	121.5	3.4	1719	3	US-08-399-411-4		Sequence 4, Appl
27	121.5	3.4	1719	3	US-08-516-859A-4		Sequence 4, Appl

Query Match	4.2%;	Score 152;	DB 2;	Length 245;
Best Local Similarity	27.4%;	Pred. No. 5.3e-05;		

286 EDVRAAAGKPKMEETLTCTICODLLHDCVSLPCMHTFCAACYSGWMERSL-CPTCRC 344

..

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Db 3 EELSAAT---SYTEDDFYCPVCQEVLTPTVTAACQHVFCRCKFLTAMRESGIHCPLCRG 59
QY 345 PV---ERICKNHIL-----NNLVEAYLIQHPDKS-----RSEEDVQSNDA 381
Db 60 SVTRERACPERALDENIMRRFSGSCRCCKKIKFYRHHYKSKYQYGVGVSSVIP 119
QY 382 RNKITQDMLQPKVRRFSDEGSSEDLLELSDVDSESSDISQPVVVCRCQPE--YRQ-- 437
Db 120 NFKISQDSV-----RSSNRSETSASDNTETQEDTSSS--GHPTFKCPLCQESNFTQRRL 172
QY 438 -----AAQPPHCP-----APEGEP 451
Db 173 LDHCNSNHLFQIVPVTCFICVSLPWGDP 200

RESULT 2

US-09-252-329-36
; Sequence 36, Application US/09252329
; Patent No. 6147192
; GENERAL INFORMATION:
; APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
; TITLE OF INVENTION: Weight Control Pathway Genes and Uses
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/252,329
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/897,340
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-005CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-252-329-36

Query Match 4.2%; Score 152; DB 3; Length 245;
Best Local Similarity 27.4%; Pred. No. 5.3e-05;
Matches 57; Conservative 22; Mismatches 77; Indels 52; Gaps 10;
QY 286 EDVRAAGKDFKMEETLTICIQDLHDCVSLQPCNHTFCAACYSYSGWMSRL-CPTCRC 344
Db 3 EELSAAT---SYTEDDFYCPVCQEVLTPTVTAACQHVFCRCKFLTAMRESGIHCPLCRG 59
QY 345 PV---ERICKNHIL-----NNLVEAYLIQHPDKS-----RSEEDVQSNDA 381
Db 60 SVTRERACPERALDENIMRRFSGSCRCCKKIKFYRHHYKSKYQYGVGVSSVIP 119
QY 382 RNKITQDMLQPKVRRFSDEGSSEDLLELSDVDSESSDISQPVVVCRCQPE--YRQ-- 437

Db 120 NFKISQDSV-----RSSNRSETSASDNTETQEDTSSS--GHPTFKCPLCQESNFTQRRL 172
QY 438 -----AAQPPHCP-----APEGEP 451
Db 173 LDHCNSNHLFQIVPVTCFICVSLPWGDP 200

RESULT 3

US-07-945-283-7
; Sequence 7, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EPO and LIT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtis P.
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext.513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 67 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Varicella-zoster virus
US-07-945-283-7

Query Match 3.8%; Score 135; DB 1; Length 67;
Best Local Similarity 36.7%; Pred. No. 0.00022;
Matches 22; Conservative 6; Mismatches 32; Indels 0; Gaps 0;

QY 303 TCICQDLHDCVSLQPCNHTFCAACYSYSGWMSRL-CPTCRCVERICKHILNLVEAY 362
Db 5 TCTICMSTVSDLGKWTMPCLHDFCVCIRAWTSTVQCPLCRCPVQSLILKHIVSDTSYKEY 64

RESULT 4

US-08-897-340-35
; Sequence 35, Application US/08897340
; Patent No. 5953306
; GENERAL INFORMATION:
; APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
; TITLE OF INVENTION: Weight Control Pathway Genes and Uses
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts

```

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,340
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/715,032
; FILING DATE: 17-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-005CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-897-340-35

Query Match 3.8%; Score 135; DB 2; Length 245;
Best Local Similarity 26.0%; Pred. No. 0.0015;
Matches 54; Conservative 23; Mismatches 79; Indels 52; Gaps 10;

QY 286 EDVRAAGKPKMEETLTCTICQDLHDCVSLQPCMHFTCAACY-SGWMERSLCPTRC 344
Db 3 EDLSAAT---SYTEDDFYCPVCQEVLTPTVTTACQHFCKFLTAMRESGAHCPLCRG 59

QY 345 PV---ERICKNHILN-----NLVEAYLIQHDPKS-----RSBEDVQSMDA 381
Db 60 NVTRERACPERALDLENIMKFGSGCRCCAKQIKFYMRHHYKSKKYQDEYGVSSIIP 119

QY 382 RNKITQDMLQPKVRRSFDSEGSSEDLLELSVDSESDISQPVVVCRCQPE--YRRQ-- 437
Db 120 NFQISQDSVGNR-----SETSDNTETQENTSSS--GHPTFKCPLQESNFTRQL 172

QY 438 -----AAQPPHCP-----APEGEF 451
Db 173 LDHCNSNHLFQIVPTCTPCVSLPWGDP 200

RESULT 5
US-09-252-329-35
; Sequence 35, Application US/09252329
; Patent No. 6147192
; GENERAL INFORMATION:
; APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
; TITLE OF INVENTION: Weight Control Pathway Genes and Uses
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/252,329
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/897,340
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-005CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 245 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-09-252-329-35

Query Match 3.8%; Score 135; DB 3; Length 245;
Best Local Similarity 26.0%; Pred. No. 0.0015;
Matches 54; Conservative 23; Mismatches 79; Indels 52; Gaps 10;

QY 286 EDVRAAGKPKMEETLTCTICQDLHDCVSLQPCMHFTCAACY-SGWMERSLCPTRC 344
Db 3 EDLSAAT---SYTEDDFYCPVCQEVLTPTVTTACQHFCKFLTAMRESGAHCPLCRG 59

QY 345 PV---ERICKNHILN-----NLVEAYLIQHDPKS-----RSBEDVQSMDA 381
Db 60 NVTRERACPERALDLENIMKFGSGCRCCAKQIKFYMRHHYKSKKYQDEYGVSSIIP 119

QY 382 RNKITQDMLQPKVRRSFDSEGSSEDLLELSVDSESDISQPVVVCRCQPE--YRRQ-- 437
Db 120 NFQISQDSVGNR-----SETSDNTETQENTSSS--GHPTFKCPLQESNFTRQL 172

QY 438 -----AAQPPHCP-----APEGEF 451
Db 173 LDHCNSNHLFQIVPTCTPCVSLPWGDP 200

RESULT 6
US-09-894-998A-47
; Sequence 47, Application US/09894998A
; Patent No. 6537555
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998A
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HSV-2
; US-09-894-998A-47

Query Match 3.8%; Score 134.5; DB 4; Length 826;
Best Local Similarity 22.0%; Pred. No. 0.01;
Matches 80; Conservative 33; Mismatches 129; Indels 121; Gaps 19;

QY 152 GAGRADP---RVPPSSATQVCPEEPQPSSTSLPTASASSTPSPAG-----RE 201
Db 6 GTSSRADPGPFPPTQPTQPA--APHAWGLMNDQNLASDSEETEVGISDDDLHRD 63
```

QY 202 RSSCGSGGGISPKGSGPSVASDEVSPASALPDRKTASFSLSLEPODQEDLEPVKKMR 261
Db 64 STSEAG-----SDDTEMFEAGLMDAAT-----PPAPFPAERQ 96
QY 262 GDGDLNLGOLLVAQPRNAQTVHEDVRAAAGKPKMEETLCIIICODLLHDCVSLQ--P 319
Db 97 SPTPADAQSC-----GGGFVGESEAEAGGGD-----VCAVCTDEIAPPLRCQSP 143
QY 320 CMHTFCAACYSGWMSRSLCPTCRCPVERICKHILNNLVEAYLIQHPKSRSEEDVQSM 379
Db 144 CLHPFCIPCKMTWPLANTCLCNTPV-----AYLIVGTASGSFTIPTIV 189
QY 380 -DARKITQDMLQPKVRRSFSDEGSSDILL-----ELLS-----412
Db 190 NDRPTRVEAEA-----AVRAGTAVDFIWTGNPRTAPRSLGHTVRLSPTPPWP 240
QY 413 DVDSESSDISQ--PYV--VCEQCE-----YARQAQPHC-PAPEGEPAQ--ALG 458
Db 241 GTDDEDDDLADGVDPFAPRPRRRGGGAGATRGTSQPAATRPAP-----PGAPRSSSG 297
QY 459 DAP 461
Db 298 GAP 300

RESULT 7

US-08-227-536-2
; Sequence 2, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
; APPLICANT: Ecken, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/227,536
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Ph.D., Kathleen A.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: DFCI-308XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2414 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-536-2

Query Match 3.7%; Score 132.5; DB 1; Length 2414;
Best Local Similarity 21.1%; Pred. No. 0.078;
Matches 128; Conservative 75; Mismatches 235; Indels 169; Gaps 31;
QY 4 PEEKQSPFPQ-----PWGRLRLGAEGBPH-----VLL-----RKREWTI 40

Db 300 PNMGOQ--PAPQVQPPGLVTPVAQMGSGAHTADPEKRKLIQQOLVLLHAHKCORREOAN 358
QY 41 GRRRGCDLSE---PSNKLVSGDHCRIVWDEKSOVLEDTSTSGTVINKLVKVVKQTCPL 97
Db 359 GEVRQCNLPHCRTMKNVNHWHCO---SGKCOVA--HCASSRQIISHWKNCTRHDCPV 413
QY 98 -----QTGDVILYVRKNEPEHNVAVLYESLSKQKMTQESFE--ANKENVFHGYKDTSG 150
Db 414 CLPLKXNAGD-----XNQOQPIILTGAPVGLGNPSLGVGQOSAPNLTVSQIDPSSIERAY 468
QY 151 AGAGRGADPRVPSSPATQVCFFEE--PQPSTSTSDLPFTASASTSTSPSPAGRRSSSCSG 209
Db 469 AALGLPQVQNMQTPQVQAKNQONQPGQSPQGMFPMNSVA---SPMG-----VNG 518
QY 210 GGGISPKGSGPSVASDEVSPASALPDR-----KTASFSLSPQD---QEDLEPVKKMR 261
Db 519 GVGQVT-----PSLSD--SMLHSAINSONPMMSSENASVPSLGFMTAAQPSSTTGIRKQWH 572
QY 262 GDGDLNLGOLLVAQPRNAQTVHEDVRAAAGKPD-----KVEETLTCI--ICODLLH 312
Db 573 EDITQDLRNHL-----VHKLVOAIFPTPDPAALKDARMENLVAYARKVEGDMYE 621
QY 313 DCVSLQPCMTTFCAACYSGWMSRSLCPTCRCPVERICKHILNN---LVEAYLIQHPDK 369
Db 622 SANRAEYHLLAEKIYKIQELEEKRT-----RLQKNMLPNAAGHVPVSMPPGNM 675
QY 370 SRSEEDVQSM-----ARNKITQDMLQPKVRRSFSDEGSSDILLSDVDESSDIS 422
Db 676 GQPPQGMTSNGPLPDPSPMIRGSPVNMMPRITPQSGLNQFG-----QMSMA 721
QY 423 QPVVCRQCP--RYRQAQAP-----PHCPAPEGE-----P 451
Db 722 QPFIIVPQTPPLQHHQQLQPGALAPPNGYGRMOPQPNQOGFLPQTQPPSGMVTNIP 781
QY 452 GAPOALGDAPSTSVLTAVQDVVCL-----QGSHALCTCCFQMPDRRAEED 502
Db 782 LAPSS--QAPVSOAQWSSS-----SCPVNSPIMPPSGSGSHIHPQLPQP-----ALHQS 831
QY 503 PRVAPQ 509
Db 832 PSPVSR 838

RESULT 8

PCT-US95-04682-2
; Sequence 2, Application PC/TUS9504682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04682
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,536
; FILING DATE: 14-April-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-308Xg999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04682-2

Query Match 3.7%; Score 132.5; DB 5; Length 2414;
Best Local Similarity 21.1%; Pred. No. 0.078;
Matches 128; Conservative 75; Mismatches 235; Indels 169; Gaps 31;
QY 4 PERGKSPPPQ-----PWGLRLILGAEGPH-----VLL-----RKREWTI 40
DB 300 PNMGOQ-PAPVQOPGLVTPVAQMGSGAHTADPEXKLIQQOVLVLLHAHKCORREQAN 358
QY 41 GRERGDLSP---PSNKLVSQDHCRIWVDEKSGQVLTEDTSTGTVINKLVVKKQTCLP 97
DB 359 GEVRCQNLPHCRTKKNVLEHWHQ---SGKSCQVA--HCASSRQIISHWKNCTRHDQCV 413
QY 98 -----QTGVDVYLVYRNKEPHEVNAVLYESLSKQMTQESPE--ANKENVFHGKDTSG 150
DB 414 CLPLKNAGD-----XRNQPIILTGAPVGLGNPSLGVGQOSAPNLSTVQIDPSSIERAY 468
QY 151 AGAGRGADPRVPSSPATQVCFEE-POPSTSTSDLPFTASASTEPSAGRERSSCSGS 209
DB 469 AALGLPYQNMTOFQVQAKNQONQPGQSPQGMPEMNSA---SPWG-----VNG 518
QY 210 GGGISPKGSPSVASDEVSSFSASALPDR-----KTASFSLPEQD---QEDLEPVKKQKR 261
DB 519 GVGVTQ-----PSLISD--SMLHSAINSQNPMSENASVPSLGPMPAAQPSITGIRKQWH 572
QY 262 GGDGLDLNGOLLVAQPRNAQTVEDVRAAGKPD-----KMEETLACI--ICODLLH 312
DB 573 EDITQDLRNLH-----VHKLVCQAITPDPDPAALKDRKRNELVYARKVGDWYE 621
QY 313 DCVSLQPCMTTCAACYSQWMSRSLCPTCRCPVERICKNHLNN---LVEAYLIQHDPK 369
DB 622 SANRABVYHLLAEKIYKIQELEKERT-----ALQKNMLPNAAGVVPVSMNPGPNM 675
QY 370 SRSEEDVQSM-----ARKITQDMLQPKVRRSFSDEGSSDLELSDVDSSESDIS 422
DB 676 GQOPQGMTSNGPLPDPMSIRGSPVNMPPRITPQSGLNQFG-----QMSMA 721
QY 423 QPVVVCRCQP---EYRQAAQF-----PHCPAPEGE-----P 451
DB 722 QPPIVPRQTPPLQHHGQLAQPGALNPMWGYGPRMQQPSNQQLPQTQPPSQGMVNTIP 781
QY 452 GAPOALGDAPSTSVSLTAVQDVVCP-----QGSALCTCCFQMPDRAERBOD 502
DB 782 LAPSS-QQAPVSOQMSSS-----SCPVNSPIMPFGSQGSHICPQLFOP-----ALHQS 831
QY 503 PRVAPQ 509
DB 832 PSPVPSR 838

US-08-296-791-6
Sequence 6, Application US/08296791
Patent No. 6245337
GENERAL INFORMATION:
APPLICANT: St. Gene III, Joseph W.
APPLICANT: Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1848 amino acids
TYPE: amino acid
TOPOLOGY: unknown
US-08-296-791-6

Query Match 3.7%; Score 131; DB 3; Length 1848;
Best Local Similarity 19.9%; Pred. No. 0.07;
Matches 99; Conservative 67; Mismatches 185; Indels 146; Gaps 22;
QY 53 NKLVSQDHCRIWVDEKSG-----QVLTEDTSTGTVINKLVVKKQTCLQTDGVYLVY 107
DB 929 NKSATGNFTLQVAD-KTGEFNHNLTLFDASNA--TRNLEVT-----LANGSVDRGAM 979
QY 108 RNKEPHEVNAV-LYESLSKQMTQEQ-----SFEANKENVFHGKDTSGA 151
DB 980 KYKLVNNGRDLNYPEVEKKNQIVDTNITPNDIQADAPSAQSNNEIAR-----1031
QY 152 GAGRGADPRVPSSPATQVCFEEPOPSTSTSD-----LFTASASTEPSAGRERSSC 206
DB 1032 -----VETPVFPAPATESALASEQPETRAPETAQAPEMETNTANSTETAPKS-----1079
QY 207 GSGGGISPKGSPSVASDEVSSFSASALPDRKTASFSLPEQDQEDLEPVKKMGDGL 266
DB 1080 -----DTATQTNPNSESVPSETTEKVAENPPQENETV--AKNEQATEPT 1123
QY 267 DLNGOLLA-QPRRNAQTVEDVRAAGKPD-----KMEETLTCIICDILLHDCVSL- 317
DB 1124 PQNGEVAKEDQPTVEANTQTEATQSEKGTETQTAETKSEPTESVTVSENQPEKTVSQS 1183
QY 318 -----QPCNHTTCAACYSQWMSRSLCP--TCRCP-----VERICKNHLNN 357
DB 1184 TEDKVVVKEEKARVET-----EETQKQVQVTSKEPPKQABPAPEVPTD--TN 1230
QY 358 LVEAYLIQH-----PKSRSEEDVQSMDEARKITQDMLQPKVRRSFSDEGS 404
DB 1231 AEEAQLQQTQPTTVAAAEFTSPNSKPAEETQQPS-----KTNAEPVTPVSENTAQTE 1287
QY 405 SEDLLESDVDSSESDISQPVVVCRCQPEYRQAAQPPHCPAPEGEPG-----452
DB 1288 TE---ETARKVEKETQ-EVPQVSAQESPKQPAKQPAQTKPOABPARENVLTKNVE 1343
QY 453 -APOALGDAPSTSVSLT 468
DB 1344 POPAQPOQTQSTAVPTT 1360

RESULT 10

US-09-839-996-6
; Sequence 6, Application US/09839996
; Patent No. 6642371
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/839,996
; FILING DATE: 20-Apr-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1848 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-839-996-6
Query Match 3.7%; Score 131; DB 4; Length 1848;
Best Local Similarity 19.9%; Pred. No. 0.07;
Matches 99; Conservative 67; Mismatches 185; Indels 146; Gaps 22;
QY 53 NKLVS GDHCRIVVDEKSG-----QVTLEDTSTSGTVINKLVKVKQTCPLOTGDIYLVY 107
Db 929 NKSATGNFTLQVAD-KTGPNNHNLTLFDASNA--TRNNLEVT-----LANGSVDRGAW 979
QY 108 RKNPEHNWAY-LYESLSSEKQGTQZ-----SFEANKENVFHGKTDTSGA 151
Db 980 KYKLRNVNGRYDLYNPEVEKRNQTVDTNITTPNDIQADAPSAQSNNEIAR----- 1031
QY 152 GAGRGADPRVPPSSPATQVCFFEPQSTSTSD-----LFFTASASSTSPSPAGRERSSC 206
Db 1032 -----VETVPPPPAPATESAIASEQETPRAETAQAQPAEETNTANSTETAPKS----- 1079
QY 207 GSGGGGISPKGSPSVASDEVSSPASALPDRKTASFSSLEPQDQDLPEVKKMKRGDGL 266
Db 1080 -----DTATQTNPNSESVSETEKVAENPPQENETV--AKNEQATEPT 1123
QY 267 DLNGQLLVA-QPRRNAQTVHEDVRAAAGKPD-----KMEETLTCTIICQDLLHDCVSL- 317
Db 1124 PQNGEVAKEDQPTVEANTQTNATQSEGKTEETQTAETKSEPTESVTVSENQPEKTVSQS 1183
QY 318 -----QPCMTTFCAACVSGMWRSSLCPP--TCRCP-----VERICKNHLN 357
Db 1184 TEDKVVVEKEEKAKVET-----EETQKAPQVTSKEPPKQABPAPEEPTD---TN 1230
QY 358 LVEAYLIQH-----PDKSRSEEDVQSM DARNKITQDMLQPKVRRSFSDERGS 404
Db 1231 ABEAQAQQOQPTTVAAAEETTSFNKPAEETQQPSE---KTNAEPVTPVVSSENTATQPT 1287
QY 405 SEDLLELSVDSSDISQPYVVCRCQCEPYRRQAQPHCHPAPEGEPG----- 452
Db 1288 TE---ETAKVEKEKQ--EVPQVASQESPKQEQPAKPAQATKQPAEPARENVLTNNVGE 1343
QY 453 -APQALGDAPSTSVSLT 468
Db 1344 PQPQAQPTQSTAVPTT 1360

RESULT 11
US-10-080-505-6
; Sequence 6, Application US/10080505
; Patent No. 6676948
; GENERAL INFORMATION:
; APPLICANT: St. Geme, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTIENS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 1848
; TYPE: PRM
; ORGANISM: Haemophilus influenzae
US-10-080-505-6
Query Match 3.7%; Score 131; DB 4; Length 1848;
Best Local Similarity 19.9%; Pred. No. 0.07;
Matches 99; Conservative 67; Mismatches 185; Indels 146; Gaps 22;
QY 53 NKLVS GDHCRIVVDEKSG-----QVTLEDTSTSGTVINKLVKVKQTCPLOTGDIYLVY 107
Db 929 NKSATGNFTLQVAD-KTGPNNHNLTLFDASNA--TRNNLEVT-----LANGSVDRGAW 979
QY 108 RKNPEHNWAY-LYESLSSEKQGTQZ-----SFEANKENVFHGKTDTSGA 151
Db 980 KYKLRNVNGRYDLYNPEVEKRNQTVDTNITTPNDIQADAPSAQSNNEIAR----- 1031
QY 152 GAGRGADPRVPPSSPATQVCFFEPQSTSTSD-----LFFTASASSTSPSPAGRERSSC 206
Db 1032 -----VETVPPPPAPATESAIASEQETPRAETAQAQPAEETNTANSTETAPKS----- 1079
QY 207 GSGGGGISPKGSPSVASDEVSSPASALPDRKTASFSSLEPQDQDLPEVKKMKRGDGL 266
Db 1080 -----DTATQTNPNSESVSETEKVAENPPQENETV--AKNEQATEPT 1123
QY 267 DLNGQLLVA-QPRRNAQTVHEDVRAAAGKPD-----KMEETLTCTIICQDLLHDCVSL- 317
Db 1124 PQNGEVAKEDQPTVEANTQTNATQSEGKTEETQTAETKSEPTESVTVSENQPEKTVSQS 1183
QY 318 -----QPCMTTFCAACVSGMWRSSLCPP--TCRCP-----VERICKNHLN 357
Db 1184 TEDKVVVEKEEKAKVET-----EETQKAPQVTSKEPPKQABPAPEEPTD---TN 1230
QY 358 LVEAYLIQH-----PDKSRSEEDVQSM DARNKITQDMLQPKVRRSFSDERGS 404
Db 1231 ABEAQAQQOQPTTVAAAEETTSFNKPAEETQQPSE---KTNAEPVTPVVSSENTATQPT 1287
QY 405 SEDLLELSVDSSDISQPYVVCRCQCEPYRRQAQPHCHPAPEGEPG----- 452
Db 1288 TE---ETAKVEKEKQ--EVPQVASQESPKQEQPAKPAQATKQPAEPARENVLTNNVGE 1343
QY 453 -APQALGDAPSTSVSLT 468
Db 1344 PQPQAQPTQSTAVPTT 1360

RESULT 14
US-08-764-870-14
; Sequence 14, Application US/08764870

```

RESULT 12
PCI-US95-10661A-6
; Sequence 6, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1848 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
;
PCI-US95-10661A-6
Query Match 3.7%; Score 131, DB 5; Length 1848;
Best Local Similarity 19.3%; Pred.No. 0.07;
Matches 99; Conservative 67; Mismatches 185; Indels 146; Gaps 22;

QY 53 NKLVSGDHCRIVVDEKSG-----QVTLEDTSTSGTVINKLVKKQTCPLOTGDIYIIVY 107
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 929 NKSATGNFTLOVAD-KTGEPNHNELTLFDASNA--TRNNLEV-----LANGSVDRGAW 979
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
QY 108 RKNPEHHWAY-LYESLSKEQGKTQE-----SFEANKENVFGTKDTSGA 151
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 980 KYKLNVNGRYDILYNVEVERKRNQVDTITTPNDIQADAPAAQSNNNEIAR----- 1031
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
QY 152 GAGRGADPRVPSPATQVCFFEPQSTSTD-----LFFTASASTESPAGRERSSSC 206
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 1032 -----VETPPVPAPATESAIASEQETRAPETAQAQPAEMEINTANSTETAPKS----- 1079
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
QY 207 GSGGGGISPKGSFVSASDEVSSFASALPDKRTASFSSLSLEPDQODELPVKKKMRGDGL 266
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 1080 -----DFATQTNENSNSSVSETEKVAENPPGENETV---AKNEQATEPT 1123
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
QY 267 DLNQQLIVA-QPRRNAQTVDHYRAAAGRPD-----RMEETLTCLIIQQDLLHDCVSL- 317
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 1124 PONGEVAKEQPVTVEANTQTNEATQSEGKTEETQTAEKSEPTESVTVENQPEKTVSOS 1183
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
QY 318 -----QPCMHFTFCACYSGWHERSSLCP--TCRCP-----VERICKHILNN 357
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db 1184 TEDKVVVVEEKAKVET-----EETQKAPQVTSKEPKQAPPAPEEPTD----TN 1230
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
QY 358 LVZEAYLIQH-----PDKRSSEEDVQSDARNKIQTDMLOPKVRKSFSDDEGS 404
   |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:

```

Patent No. 6236946
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S
APPLICANT: Baxter, John D
APPLICANT: Fletterick, Robert J
APPLICANT: Wagner, Richard L
APPLICANT: Kushner, Peter J
APPLICANT: Apriletti, James W
APPLICANT: West, Brian
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: Cooley Godward
STREET: Five Palo Alto Square, 3000 El Camino Real
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,870
FILING DATE: 13-DEC-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,540
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,543
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,606
FILING DATE: 14-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Jackie N
REGISTRATION NUMBER: 35,966
REFERENCE/POCKET NUMBER: UCAL-246/01US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)843-5000
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 933 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-764-870-14

Query Match 3.5%; Score 126.5; DB 3; Length 933;
Best Local Similarity 21.3%; Pred. No. 0.061;
Matches 117; Conservative 60; Mismatches 196; Indels 177; Gaps 30;
QY 124 SEKQMTQESFEANKENVFHGTDTSGAGAGRADPR-----VPPSSPATQV 170
DB 70 SDEKTDQOQSL-SDVEGAYSRAEATRGAGSSSPPEKDSGLDLSVLDTLAPSGFG--- 125
QY 171 CFEPQPS-----TSTSDLF-----PTASASSTEPSFAGRRSSSCGSGGG----- 211
DB 126 --SQSPSPACEVTSSWCLFGPELPDPAATQVLSPLMSRSGCKVGDSSGTAATAH 182
QY 212 GISPKGSGPS-----VASDEVSSFASALPRKKTASFSSLEPQDQEDLEP-----VKKK 259
DB 183 KVLPRGLSPARQLLLPASE--SPHWSGAPVKPSPQAAAVEEEDSESEESAGPLLKGR 240
QY 260 MRGDDGLDLNGLLVAOPRANAQTV-----HEDVRAAAGKPDKNEE-----TLTICIC 307
DB 241 PRALGGAAGGGAACFPGAAGGVALVPKEDSRFSAPRVALVEQDAPMAPGRSPLATTV 300
QY 308 QDLHLDCVSLQPCXHTFCAA-----CYSQWM-----ERSSLC----- 339

DB 301 MDTFH--VPILPLNHALLAARTQLEDESVDGAGASAFAPPTSPCASSTPVAVGDF 358
QY 340 PTCRCFVERICKNHILNNLVEAYLI-----QHPD-KRSEEDVQSMARKIT----- 386
DB 359 PDCAYPDPAEPKD-----DAYPLYSDFQPPALKIKEEEGAGASARSFVYVAGANP 411
QY 387 -----QDMLQPKVRSPSDESG--SSDLELLESDVDSESDISQFVVCRCQPEY 434
DB 412 AAFPDPFLGPPLPPRATPSRFGAAVTAAPASASVSSASSGSTL-----ECILY 463
QY 435 RRQAAQP---PHCPAPEGEPGAPQAL---GDAPSTSVSLTTA-VQYVVCFLQGSALCTC 487
DB 464 KAEGAPQOGPEAPPCKAPGASGCLLPDGLPSTSSASAAAAGAAPALYPALGLNGLPOL 523
QY 488 CFQP-----MPDRAAREQDPRA-----PQQ-CAVCLQPFCHLYWGC 524
DB 524 GYQAAVLKEGLQVYPPYLYNLYRPD--SEASQPSQVSESLFKICLIC-----GD 572
QY 525 TRTGC-YGCL 533
DB 573 EASGCHYGLV 582
RESULT 15
US-08-980-115-14
Sequence 14, Application US/08980115
Patent No. 6266622
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S.
APPLICANT: Baxter, John D.
APPLICANT: Fletterick, Robert J.
APPLICANT: Wagner, Richard L.
APPLICANT: Kushner, Peter J.
APPLICANT: Apriletti, James W.
APPLICANT: West, Brian L.
APPLICANT: Shiau, Andrew K.
TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
CURRENT APPLICATION NUMBER: US/08/980,115
CURRENT FILING DATE: 1997-11-26/08/764,870
EARLIER FILING DATE: 1996-12-13
EARLIER APPLICATION NUMBER: 60/008,606
EARLIER FILING DATE: 1995-12-14
EARLIER APPLICATION NUMBER: 60/008,543
EARLIER FILING DATE: 1995-12-13
EARLIER APPLICATION NUMBER: 60/008,540
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 933
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (659)..(918)
OTHER INFORMATION: minimal ligand binding domain
US-08-980-115-14

Query Match 3.5%; Score 126.5; DB 3; Length 933;
Best Local Similarity 21.3%; Pred. No. 0.061;
Matches 117; Conservative 60; Mismatches 196; Indels 177; Gaps 30;
QY 124 SEKQMTQESFEANKENVFHGTDTSGAGAGRADPR-----VPPSSPATQV 170
DB 70 SDEKTDQOQSL-SDVEGAYSRAEATRGAGSSSPPEKDSGLDLSVLDTLAPSGFG--- 125
QY 171 CFEPQPS-----TSTSDLF-----PTASASSTEPSFAGRRSSSCGSGGG----- 211
DB 126 ---SQSPSPACEVTSSWCLFGPELPDPAATQVLSPLMSRSGCKVGDSSGTAATAH 182

Search completed: May 7, 2004, 14:51:21
Job time : 46.1732 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:49:13 ; Search time 105.174 Seconds
(without alignments)
1752.371 Million cell updates/sec

Title: US-10-048-046-2

Perfect score: 3585
Sequence: 1 MERPEEGKQSPPPQWGRLL.....VKAHAKFNHICQTRFKN 664

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	3585	100.0	664	9	US-09-780-525-2
2	1954.5	54.5	426	9	US-09-764-864-809
3	1803	50.3	349	15	US-10-108-260A-4516
4	843.5	23.5	120	9	US-09-764-864-1270
5	566	15.8	238	15	US-10-094-749-1799
6	474	13.2	92	9	US-09-764-864-1109
7	432	12.1	99	9	US-09-764-864-1110
8	432	12.1	99	9	US-09-764-864-1529
9	252	7.0	287	12	US-10-424-599-264601
10	146.5	4.1	284	12	US-10-126-103-114
11	146	4.1	303	15	US-10-108-260A-3965
12	142	4.0	648	12	US-10-425-114-63379
13	139	3.9	775	14	US-10-224-999A-3462
14	139	3.9	825	14	US-10-121-988-161
15	139	3.9	825	14	US-10-200-562-161

16	139	3.9	825	14	US-10-237-551-161	Sequence 161, App
17	139	3.9	825	14	US-10-210-428-1	Sequence 1, Appli
18	139	3.9	4675	15	US-10-093-463-74	Sequence 74, Appl
19	139	3.9	4691	15	US-10-093-463-72	Sequence 72, Appl
20	138	3.8	255	14	US-10-106-698-4393	Sequence 4393, Ap
21	138	3.8	376	12	US-10-424-599-233512	Sequence 233512,
22	136.5	3.8	420	12	US-10-425-114-70107	Sequence 10707, A
23	136.5	3.8	424	12	US-10-425-114-70107	Sequence 118, App
24	136	3.8	521	15	US-10-259-194A-118	Sequence 9, Appli
25	135	3.8	245	9	US-09-998-667-9	Sequence 841, App
26	135	3.8	285	9	US-09-764-864-841	Sequence 3125, Ap
27	135	3.8	285	15	US-10-364-049-3125	Sequence 56089, A
28	135	3.8	478	12	US-10-425-114-56089	Sequence 72, Appl
29	135	3.8	993	15	US-10-259-194A-72	Sequence 187336,
30	134.5	3.8	353	12	US-10-424-599-187336	Sequence 47, Appl
31	134.5	3.8	826	14	US-10-121-988-47	Sequence 47, Appl
32	134.5	3.8	826	14	US-10-200-562-47	Sequence 47, Appl
33	134.5	3.8	826	14	US-10-237-551-47	Sequence 37376, A
34	134.5	3.8	826	14	US-10-425-114-37376	Sequence 2672, Ap
35	134	3.7	331	12	US-10-104-047-2672	Sequence 65516, A
36	133.5	3.7	610	15	US-10-425-114-65516	Sequence 54084, A
37	133	3.7	346	12	US-10-425-114-54084	Sequence 203186,
38	132.5	3.7	440	12	US-10-424-599-203186	Sequence 5784, Ap
39	132.5	3.7	441	12	US-10-369-493-5784	Sequence 818, App
40	132.5	3.7	3507	15	US-09-764-864-818	Sequence 387, App
41	132	3.7	658	9	US-09-764-864-818	Sequence 1277, Ap
42	131.5	3.7	191	16	US-10-389-566-387	Sequence 6, Appli
43	131.5	3.7	563	9	US-09-764-864-1277	Sequence 6, Appli
44	131	3.7	1848	10	US-09-839-996-6	Sequence 6, Appli
45	131	3.7	1848	12	US-10-645-655-6	

ALIGNMENTS

RESULT 1

US-09-780-525-2
; Sequence 2, Application US/09780525
; Patent No. US20020004223A1
; GENERAL INFORMATION:
; APPLICANT: Bin-Bing Zhou
; APPLICANT: Yuan Zhu
; APPLICANT: Priya Chaturvedi
; APPLICANT: Mark R. Hurler
; APPLICANT: Xiaotong Li
; TITLE OF INVENTION: FHAR1, A NEW RING FINGER PROTEIN
; FILE REFERENCE: GP-70668-C1
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US/09/780,525
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 664
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-780-525-2

Query Match	100.0%;	Score	3585;	DB	9;	Length	664;
Best Local Similarity	100.0%;	Pred. No.	5e-271;				
Matches	664;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MERPEEGKQSPPPQWGRLLRLGAE	GEHPHVLRLKREWTIGRRGCDLSFP	SNKLVSGDH	60		
Db	1	MERPEEGKQSPPPQWGRLLRLGAE	GEHPHVLRLKREWTIGRRGCDLSFP	SNKLVSGDH	60		
QY	61	CRIVVDEKSGQVLTEDTSGTGVINKL	KVKKQTPLQGDVILVYRKNEPHNAVLY	120			
Db	61	CRIVVDEKSGQVLTEDTSGTGVINKL	KVKKQTPLQGDVILVYRKNEPHNAVLY	120			
QY	121	ESLSERKQGTQBSFRANKENVFHGT	YDTSAGAGRGADPRVPPSPATVCP	EEPOPSTS	180		

Db 121 ESLSEKQGMTOESFEANKENVFHGTGKDTSGAGAGADPRVPPSSPATQVCFEPPQSTS 180
Qy 181 TSDLPPTASASTPSPAGRRSSSCGSGGGISPKGSGPSVASDEVSSPASALPDRKTA 240
Db 181 TSDLPPTASASTPSPAGRRSSSCGSGGGISPKGSGPSVASDEVSSPASALPDRKTA 240
Qy 241 SFSSLEPOQDELEPVKKKMGDGLDLNGQLLVAQPRNAQTVHEDVRAAAGKPDQVEE 300
Db 241 SFSSLEPOQDELEPVKKKMGDGLDLNGQLLVAQPRNAQTVHEDVRAAAGKPDQVEE 300
Qy 301 TLTCTIICODLLHDCVSLQPCMHFTCAACYSWMMERSSLCPTCRCPVERICKNHIILNLVE 360
Db 301 TLTCTIICODLLHDCVSLQPCMHFTCAACYSWMMERSSLCPTCRCPVERICKNHIILNLVE 360
Qy 361 AYLIOHPDKSRSEEDVQSMADARNKITQDMLQPKVRRSFSDEEGSSDLELSDVDSESSD 420
Db 361 AYLIOHPDKSRSEEDVQSMADARNKITQDMLQPKVRRSFSDEEGSSDLELSDVDSESSD 420
Qy 421 ISQPVVCRQCEYRRAQAQPPHCPAPGEPGAPQALGDAPSTSVSLTTAVQDYVCPLOG 480
Db 421 ISQPVVCRQCEYRRAQAQPPHCPAPGEPGAPQALGDAPSTSVSLTTAVQDYVCPLOG 480
Qy 481 SHALCTCCFPMPDRRAEREDQPRVAPQCAVCLQPPCHLYWGCTRTGCGYGLAPFCFLN 540
Db 481 SHALCTCCFPMPDRRAEREDQPRVAPQCAVCLQPPCHLYWGCTRTGCGYGLAPFCFLN 540
Qy 541 LGDKCLDGLVNNNSYESDILKXLYLATRGLTWKXNMLTSLVALQRGVFLLSYRVGTDTVL 600
Db 541 LGDKCLDGLVNNNSYESDILKXLYLATRGLTWKXNMLTSLVALQRGVFLLSYRVGTDTVL 600
Qy 601 CYCCGLRSFRELTYQYRONIPASELPVAVTSRPPCYGNCRTQVKAHAMKFNHICEQT 660
Db 601 CYCCGLRSFRELTYQYRONIPASELPVAVTSRPPCYGNCRTQVKAHAMKFNHICEQT 660
Qy 661 RFKN 664
Db 661 RFKN 664

RESULT 2
US-09-764-864-809
; Sequence 809, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 809
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (414)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (415)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (420)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-809
Query Match 54.5%; Score 1954.5; DB 9; Length 426;
Best Local Similarity 74.1%; Pred. No. 5.9e-144;
Matches 389; Conservative 5; Mismatches 22; Indels 109; Gaps 5;

Qy 1 MERPEEGKQSPPPQWGRLLRLGAEGERPHVLLRREWTIGRRRCGLDLSFFSNKLVSGDH 60

Db 3 MERPEEGKQSPPPQWGRLLRLGAEGERPHVLLRREWTIGRRRCGLDLSFFSNKLVSGDH 62
Qy 61 CRIVDEKSGQVTLTSTGTVINKLVKKQKOTCPLQTDGVIYLVYKNEPEHNVAIY 120
Db 63 CRIVDEKSGQVTLTSTGTVINKLVKKQKOTCPLQTDGVIYLVYKNEPEHR----- 117
Qy 121 ESLSEKQGMTOESFEANKENVFHGTGKDTSGAGAGADPRVPPSSPATQVCFEPPQSTS 180
Db 118 ----- 117
Qy 181 TSDLPPTASASTPSPAGRRSSSCGSGGGISPKGSGPSVASDEVSSPASALPDRKTA 240
Db 118 -----SGGGGISPKGSGPSVASDEVSSPASALPDRKTA 150
Qy 241 SFSSLEPOQDELEPVKKKMGDGLDLNGQLLVAQPRNAQTVHEDVRAAAGKPDQVEE 300
Db 151 SFSSLEPOQDELEPVKKKMGDGLDLNGQLLVAQPRNAQTVHEDVRAAAGKPDQVEE 210
Qy 301 TLTCTIICODLLHDCVSLQPCMHFTCAACYSWMMERSSLCPTCRCPVERICKNHIILNLVE 360
Db 211 TLTCTIICODLLHDCVSLQPCMHFTCAACYSWMMERSSLCPTCRCPVERICKNHIILNLVE 270
Qy 361 AYLIOHPDKSRSEEDVQSMADARNKITQDMLQPKVRRSFSDEEGSSDLELSDVDSESSD 420
Db 271 AYLIOHPDKSRSEEDVQSMADARNKITQDMLQPKVRRSFSDEEGSSDLELSDVDSESSD 330
Qy 421 ISQPVVCRQCEYRRAQAQPPHCPAPGEPGAPQALGDAPSTSVSLTTAVQDYVCPLOG 480
Db 331 ISQPVVCRQCEYRRAQAQPPHCPAPGEPGAPQALGDAPSTSVSLTT-----VRI 382
Qy 481 SHALC-----TCCQPMDRRAEREDQPRVAPQ-----QCAVCLQPP 517
Db 383 TCALCKEATPC-APAASAHARFESGTRTGPRVXXLNKCKXGLLOAF 426

RESULT 3
US-10-108-260A-4516
; Sequence 4516, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4516
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4516
Query Match 50.3%; Score 1803; DB 15; Length 349;
Best Local Similarity 99.4%; Pred. No. 3e-132; 2; Indels 0; Gaps 0;
Matches 331; Conservative 0; Mismatches 2;
Qy 205 SCGSGGGISPKGSGPSVASDEVSSPASALPDRKTASFSSLEPQDELEPVKKKMGD 264
Db 5 SFSGSGGGISPKGSGPSVASDEVSSPASALPDRKTASFSSLEPQDELEPVKKKMGD 64
Qy 265 DLDLNGQLLVAQPRNAQTVHEDVRAAAGKPDQMEETLTCIICODLLHDCVSLQPCMHFT 324
Db 65 DLDLNGQLLVAQPRNAQTVHEDVRAAAGKPDQMEETLTCIICODLLHDCVSLQPCMHFT 124
Qy 325 CAACYSWMMERSSLCPTCRCPVERICKNHIILNLVEAYLIQHPDKSRSEEDVQSMADARNK 384
Db 125 CAACYSWMMERSSLCPTCRCPVERICKNHIILNLVEAYLIQHPDKSRSEEDVQSMADARNK 184
Qy 385 ITQDMLQPKVRRSFSDEEGSSDLELSDVDSESSDISQPVVCRQCPYRRAQAQPPHC 444
Db 185 ITQDMLQPKVRRSFSDEEGSSDLELSDVDSESSDISQPVVCRQCPYRRAQAQPPHC 244

ORGANISM: Homo sapiens
FEATURE: SITE
NAME/KEY: (80)
LOCATION: (80)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1109

Query Match 13.2%; Score 474; DB 9; Length 92;
Best Local Similarity 98.9%; Pred. No. 2.9e-29;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MERPEEGKQSPPPQWGLRLRLGAECEPHVLLKREWTIGRRGCDLSFPPSNKLVSGDH 60
DB 3 MERPEEGKQSPPPQWGLRLRLGAECEPHVLLKREWTIGRRGCDLSFPPSNKLVSGDH 62
QY 61 CRIVVDEKSGQVLTEDTSGTGVINKLVV 90
DB 63 CRIVVDEKSGQVLTEDTSGTGVINKLVV 92

RESULT 7
US-09-764-864-1110
Sequence 1110, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1110
LENGTH: 99
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: SITE
NAME/KEY: (17)
LOCATION: (17)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: (96)
LOCATION: (96)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1110

Query Match 12.1%; Score 432; DB 9; Length 99;
Best Local Similarity 92.5%; Pred. No. 5.9e-26;
Matches 86; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
QY 42 RRRGCDLSFPPSNKLVSGDHCRIVVDEKSGQVLTEDTSGTGVINKLVVKKQTCPLOTGD 101
DB 9 RDKGPD--FPXNKLVSQDHCRIVVDEKSGQVLTEDTSGTGVINKLVVKKQTCPLOTGD 66
QY 102 VYLVYRKNEPEHNVAYLVESLSEKQGMTOESP 134
DB 67 VYLVYRKNEPEHNVAYLVESLSEKQGMTOESP 99

RESULT 8
US-09-764-864-1529
Sequence 1529, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1529

LENGTH: 99
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: SITE
NAME/KEY: (17)
LOCATION: (17)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: (96)
LOCATION: (96)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1529

Query Match 12.1%; Score 432; DB 9; Length 99;
Best Local Similarity 92.5%; Pred. No. 5.9e-26;
Matches 86; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
QY 42 RRRGCDLSFPPSNKLVSGDHCRIVVDEKSGQVLTEDTSGTGVINKLVVKKQTCPLOTGD 101
DB 9 RDKGPD--FPXNKLVSQDHCRIVVDEKSGQVLTEDTSGTGVINKLVVKKQTCPLOTGD 66
QY 102 VYLVYRKNEPEHNVAYLVESLSEKQGMTOESP 134
DB 67 VYLVYRKNEPEHNVAYLVESLSEKQGMTOESP 99

RESULT 9
US-10-424-599-264601
Sequence 264601, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 264601
LENGTH: 287
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_80956C.1.pep
US-10-424-599-264601

Query Match 7.0%; Score 252; DB 12; Length 287;
Best Local Similarity 24.2%; Pred. No. 3.1e-11;
Matches 83; Conservative 39; Mismatches 145; Indels 76; Gaps 10;
QY 340 PTCRCPCVERICXNHLNVLVEAYLIQHPDKSRSEEDVQSDARNKITQDMLQPKVRSFS 399
DB 2 PQCRVVFAGKNHFLRTAEDMLRADSSLSQSHDEVALDT-----YALVRSNLV 52
QY 400 DEEGSSDLELSDVDSSESDISQPVVCRQCQPEYRROAOPPHCPAGEPEGPAQALGD 459
DB 53 IGSQKKNKRAYTPLDQSDG---TYHQCCQ-----81
QY 460 APSTSVSLTAVQDYVCPQLQSGSHALCTCFQFMPDRRAERQDPRVAPQCAVCLQPFCH 519
DB 82 -----VTEVAGFRCKYDTVHLQCAQCGGMPERTGFG-----IPQYCSGCDRSPCG 127
QY 520 LYW---GCTRTGCGCLAPFCFLN-----LGDKCLDGV-----LNNSVESDILKYLATRG 568
DB 128 AYHWALGVTVNGSY-----PVCSDTLRPFISDRHSIRIPLLAHEKNLHEQNITDSCIRQMG 183
QY 569 LTWKNMLTSLVALQ-----RGVFLLSDYRVGTDLVLCYCCGLRSFRLTYQYRONIPA 622
DB 184 RTLPDVISEWIAKFENREIDRRRMLNAEMITARTFVCQDCYHKLVSFLLYWFLRLSIPK 243
QY 623 SELPVAVTSPDCYWGRCNCRTOVKA-HHAMFENHICEQTRFKN 664

[illegible][illegible]

```

RESULT 11
US-10-108-260A-3965
; Sequence 3965, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3965
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3965
Query Match 4.1%; Score 146; DB 15; Length 303;
Best Local Similarity 22.4%; Pred. No. 0.0064;

```

```

234 Db      :::|||LMURCPDPS:||||:CAAAVGQDM:-----252
405 QY SEDLLESDVDSESDISQPVVCRQCPEVYRQAAQPHCPGEPGPAQALGDAPSTS464
      :||:|:||||:|:
253 Db  NSLADVDETEKGYRLRSYI-----EDNRKT--KMCAPGCCYAAEFVMGSG-----298
      :||:|:||||:|:
465 QY VSLTAVTDYVCPQLGGSHALCTCFQPMPPRAEREQDPRA-----506
      :||:|:||||:|:
299 Db  ---SYDVNCNCSYGCWNCTE-----EAHRPVDCATVSKWILKNSAESNNMWIL345
      :||:|:||||:|:
507 QY ---PQCAVCLQPFCHLYWGTRTCY--GCLAPCELNIG-----DKCLDGLVNNNSYSD558
      :||:|:||||:|:
346 Db  ANSKCPKCKRPI-EKQGCWHITCTPFCFEFCWLCGLPWSHGERTGGFACNRYES-403
      :||:|:||||:|:
559 QY ILKNYLA-TRGLTWKNMLTESLVALQGVFLLSDYR593
      :||:|:||||:|:
404 Db  -----ARQEGVYDESRR416

```

```

RESULT 13
US-10-224-999A-3462
; Sequence 3462, Application US/10224999A
; Publication NO. US2003017318A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Composition and Method for Treating Viral Infection
; FILE REFERENCE: 5004.01
; CURRENT APPLICATION NUMBER: US/10/224,999A
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/313,695
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 3484
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3462
; LENGTH: 775
; TYPE: PRT
; ORGANISM: human herpesvirus 1
US-10-224-999A-3462

```

```

Query Match      3.9%; Score 139, DB 14; Length 775;
Best Local Similarity 22.6%; Pred. No. 0.088;
Matches      80; Conservative 34; Mismatches 142; Indels 98; Gaps 16

156 GADPRVPPSSPATQVCFEPOPST-----STSDLPPTASSTPSPAGRERSSSCGSGG 211
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6 GASTRRPEGRPO-----REPAPDVWVPCDRDL--PDSSEATEVEGR----- 48

212 GISPKGSGFSVASDEVSFASALPDRKTASFSSLEPOQEDLBFVKKKMRGGDLDLNGQ 271
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      49 -----GDADHDD-----DSASEADSTDTELFETGLLGPQVD--GGA 84

272 LLVAQPRRNAQTVHEDVRAAGKPKD-----MEETLTCLIIQDILL--HDCVSLQPCMHFFC 325
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      85 VSGGSPPR-----EEDPGSCGAPPREDGGSDGDCVACTDEIAPHLRCDTFPCMHREFC 139

326 AACYSQWMESSLCTCPCPVERICKNHI LNLVLAEAYLIQHPKSR--SEEDVQSMDAEN 383
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      140 IPCMKTWQLRNCTCLCNAGLYLLIVGTPSGSSTPIVNDQPTWMEAEAVRAGTAVD 199

384 KI-----TQDMLQPKVRRFSFDEGSSDLELSDVSESSDISQPV 426
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      200 FIWTGNQRPAPRYLLTGHTVRLSPTHPEPTTDED-----DDLLDDADYVPP-- 247

427 VCRQCEY--RRQAQPP-----HCPAPGEPCAPQA--LGDAPSTSVSLTT 469

248 APRTETRAPRERGAAPVVTGASHAAPQAPAAATAPPSAPFIPGHSSSNTTTT 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14

```

US-10-121-988-161
; Sequence 161, Application US/10121988
; Publication NO. US20030069327A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121,988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 825
; TYPE: PRT
; ORGANISM: HSV2
US-10-121-988-161

```

Query Match	3.9%;	Score	139;	DB	14;	Length	825;
Best Local Similarity	22.7%;	Pred. NO.	0.096;				
Matches	82;	Conservative	33;	Mismatches	127;	Indels	120;
						Gaps	19;
QY	152	GAGRGADP--	--KVPSPATQVCFEFPQ	STSTSDLFTTASASTPSPAG-----	RE	201	
Db	6	GTSSRADPGEPPRPRT	PGTQPA--	APHAWGLMDQWLASSDEBET	VGISDDDLHRD	63	
QY	202	RSSCSGGGGISFK	GGSPSVASDEVSS	FASALPDRKTASFSSLE	PDQDELPVKKMR	261	
Db	64	STSEAG-----	-----	STDTMEFAGLMDAAT-----	PPAPPPAERQG	96	
QY	262	GGGDLNLGQLLVAQ	PRRNaQTVEDVRAAGK	PKMEETLTCTICQDLLHDCVSLQ--	P	319	
Db	97	SPTPDAQSC-----	-----	GGGPGVEEAEAGGGD-----	VCAVCTDETA	PLRCQSP	143
QY	320	CMHTFCAACYSGM	WSSLSCTPCRCPVER	ICKNHILNNLVEAYL	IQHPDKSRSEEDVQSM	379	
Db	144	CLHPFCIPC	KTWIPLURNTCPLCNTFV-----	AYLIVGVTASGSFSTIPV	189		
QY	380	-DARNKITQDML-----	-----	QPKVR-RSFS-----	DEEGSSEDULEL	411	
Db	190	NDPRTREAEAAV	RAGTAVDFITWGNPRTAP	RSLSLGGHTVTRALS	TFPWFGTDDDEDDL	249	
QY	412	SDVDSSESDISQYV--	--VCRQCE-----	--YRQAAQPPHC--	PAPEGEFGAQ--	ALGD	459
Db	250	ADVD-----	-----	YVPFAPRPRGGG	GAGATGTSQPAATPAP--	FGAPRSSSGG	297
QY	460	AP	461				
Db	298	AP	299				

```

RESULT 15
US-10-200-562-161
; Sequence 161, Application US/10200562
; Publication No. US20030165819A1
; GENERAL INFORMATION:
; APPLICANT: McGowen, Patrick
; APPLICANT: Hosken, Nancy A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C2
; CURRENT APPLICATION NUMBER: US/10/200,562
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 825

```

```

; TYPE: PRT
; ORGANISM: HSV2
US-10-200-562-161

Query Match      3.9%; Score 139; DB 14; Length 825;
Best Local Similarity 22.7%; Pred. No. 0.096;
Matches 82; Conservative 33; Mismatches 127; Indels 120; Gaps 19;

QY 152 GAGRGADP---RVPPSSPATQVCFEEPPQSTSTSLFPTASASTESPAG-----RE 201
Db 152 GAGRGADP---RVPPSSPATQVCFEEPPQSTSTSLFPTASASTESPAG-----RE 201
QY 6 GTSSRADPGFPPRPTGTQTPA--APHAWGLNDQWMLASSDSEETEVEVGLSDDDLHRD 63
Db 6 GTSSRADPGFPPRPTGTQTPA--APHAWGLNDQWMLASSDSEETEVEVGLSDDDLHRD 63
QY 202 RSSSCSGGGGSPKSGSPSVASDEVSSPASLPDKTASPSLSBPQDQEDLEPVKKQWR 261
Db 202 RSSSCSGGGGSPKSGSPSVASDEVSSPASLPDKTASPSLSBPQDQEDLEPVKKQWR 261
QY 64 STSEAG-----STDTEMEAGLMDAAT-----PPARPPAERQG 96
Db 64 STSEAG-----STDTEMEAGLMDAAT-----PPARPPAERQG 96
QY 262 GGDGLDLNQLLVQAQPRRNAQTVHEDVRAAAGKPKDMEETLTCIQDILLHDCVSLQ--P 319
Db 262 GGDGLDLNQLLVQAQPRRNAQTVHEDVRAAAGKPKDMEETLTCIQDILLHDCVSLQ--P 319
QY 97 SPTPADAQGSC-----GGGFPVGEAEAEAGGGD-----VCAVCTDEIAPPLRCQSF 143
Db 97 SPTPADAQGSC-----GGGFPVGEAEAEAGGGD-----VCAVCTDEIAPPLRCQSF 143
QY 320 CMHTFCAACYGHWERSLLCPTCRCPVERICKNHLNLLVEAYLIQHDPKRSRSEDVQSM 379
Db 320 CMHTFCAACYGHWERSLLCPTCRCPVERICKNHLNLLVEAYLIQHDPKRSRSEDVQSM 379
QY 144 CLHFFCIPCWKTIPLRNTCPLCNTPV-----AYLIVGVTASGSFSTIPV 189
Db 144 CLHFFCIPCWKTIPLRNTCPLCNTPV-----AYLIVGVTASGSFSTIPV 189
QY 380 -DARKKITQDML-----QPKVR-RSFS-----DEEGSSDLLEL 411
Db 380 -DARKKITQDML-----QPKVR-RSFS-----DEEGSSDLLEL 411
QY 190 NDRTRVEAEAAVRAGTAVDFIWTGNPRTAPRSLSLGCHTVRALSGTPPWPFGTDDDDDL 249
Db 190 NDRTRVEAEAAVRAGTAVDFIWTGNPRTAPRSLSLGCHTVRALSGTPPWPFGTDDDDDL 249
QY 412 SDVDSSESDISQPYV--VCRQPE-----YRQAAQPPHC--PAPEGEPPAQ--ALGD 459
Db 412 SDVDSSESDISQPYV--VCRQPE-----YRQAAQPPHC--PAPEGEPPAQ--ALGD 459
QY 250 ADVD-----YVPPAPRRAPRGGGAGATRGTSQPAATRPAP---PGAPRSSSGG 297
Db 250 ADVD-----YVPPAPRRAPRGGGAGATRGTSQPAATRPAP---PGAPRSSSGG 297
QY 460 AP 461
Db 298 AP 299
```

Search completed: May 7, 2004, 15:06:49
Job time : 107.174 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:41:17 ; Search time 37.1616 Seconds
(without alignments)
1718.743 Million cell updates/sec

Title: US-10-048-046-2

Perfect score: 3585

Sequence: 1 MERPEEGKQSPFPQWGRLL.....VKAHAKFNHNCIEQTRFKN 664

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: Piri.*

2: Piri2.*

3: Piri3.*

4: Piri4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	46.8	306	T46399	hypothetical prote
2	357.5	10.0	473	C96516	F16N3.15 [imported
3	168.5	4.7	467	1 WZB861	gene 61 protein -
4	145	4.0	1603	2 S17983	gene posterior sex
5	144.5	4.0	2322	2 T10542	hypothetical prote
6	142	4.0	532	2 T49467	related to COP1-in
7	141.5	3.9	424	2 G96742	unknown protein p1
8	141.5	3.9	1208	2 T00362	hypothetical prote
9	139	3.9	775	1 EDBR11	immediate-early pr
10	139	3.9	825	1 EDBR11	immediate-early pr
11	138	3.8	387	2 T39653	probable DNA repai
12	138	3.8	639	2 T16648	hypothetical prote
13	137.5	3.8	973	2 B89009	protein T27C4.4 [i
14	136.5	3.8	1331	2 T04938	hypothetical prote
15	136.5	3.8	1495	2 A85240	hypothetical prote
16	136.5	3.8	1495	2 T10649	hypothetical prote
17	136	3.8	315	2 D48560	immediate-early pr
18	132.5	3.7	2414	2 A54277	transcription adap
19	132.5	3.7	3507	2 T34513	hypothetical prote
20	131.5	3.7	551	2 JC7562	glioblastoma RING
21	131	3.7	794	2 S59069	Z13 protein - mous
22	131	3.7	1849	2 C41859	IGA-specific metal
23	130.5	3.6	248	2 D96835	RING-H2 finger pro
24	129.5	3.6	453	2 G96895	hypothetical prote
25	129.5	3.6	776	2 T20738	hypothetical prote
26	129.5	3.6	933	1 QRHUP	progesterone recep
27	129.5	3.6	1083	2 C88854	protein F1A10.3 [
28	129	3.6	222	2 JC4296	ring finger protei
29	129	3.6	638	2 JC7753	ring finger B-box

30	129	3.6	1280	2 T00365	hypothetical prote
31	128.5	3.6	1888	2 T14273	zinc finger protei
32	127.5	3.6	245	2 T45652	RNA binding-like p
33	127	3.6	572	2 S21325	probable exo-gluca
34	126.5	3.5	406	2 S59296	probable finger pr
35	126.5	3.5	676	1 EDBR23	immediate-early pr
36	125.5	3.5	568	2 F71614	chromatinic RING f
37	125.5	3.5	989	2 A54505	serine-repeat anti
38	125.5	3.5	2163	2 T15276	hypothetical prote
39	125	3.5	630	2 A49656	estrogen-responsiv
40	124.5	3.5	1576	2 S65774	homeotic protein H
41	124.5	3.5	3869	2 A48205	All-1 protein +GRE
42	124	3.5	676	1 EDBR22	immediate-early pr
43	124	3.5	1712	2 A38261	masking protein pr
44	123.5	3.4	892	2 T09193	ataxin 7 - human
45	123.5	3.4	1233	2 T30989	serine/threonine p

ALIGNMENTS

RESULT 1

T46399

hypothetical protein DKFzp434N2420.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46399

R:Ottewaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23031

A:Accession: T46399

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-306 <AAA>

A:Cross-references: EMBL:AL137561

A:Experimental source: adult testis; clone DKFzp434N2420

C:Genetics:

A>Note: DKFzp434N2420.1

Query Match	46.8%	Score	1678	DB 2	Length	306
Best Local Similarity	100.0%	Pred. No.	1.9e-100			
Matches	306	Conservative	0	Mismatches	0	Indels
						Gaps
						0
Qy	359	VEAYLIQHPDKSRSEEDVQSDARNKITQDMLOPKVRRSFSDSEGSSEDLLESDVDS	418			
Db	1	VEAYLIQHPDKSRSEEDVQSDARNKITQDMLOPKVRRSFSDSEGSSEDLLESDVDS	60			
Qy	419	SDISQPVVVCRCQPEYRRQAQPPHCPAPEGAPQALGDAPSTSVSLTTAVQDYVCP	478			
Db	61	SDISQPVVVCRCQPEYRRQAQPPHCPAPEGAPQALGDAPSTSVSLTTAVQDYVCP	120			
Qy	479	QGSALCTCCFQPMPPRRAREQDPRVAQQAQVCLQPFCHLYWGCTRTGCGCLAPFCE	538			
Db	121	QGSALCTCCFQPMPPRRAREQDPRVAQQAQVCLQPFCHLYWGCTRTGCGCLAPFCE	180			
Qy	539	LNLDGKCLDGLVNNNSYESILKNYLATRLGKTNMLTSLVALQGVFLLSDYRTVGT	598			
Db	181	LNLDGKCLDGLVNNNSYESILKNYLATRLGKTNMLTSLVALQGVFLLSDYRTVGT	240			
Qy	599	VLCCGGLSFRLLTYQYQNTNPASLPVAVTSRDPDQYGNCRQVKAHAKFNHICE	658			
Db	241	VLCCGGLSFRLLTYQYQNTNPASLPVAVTSRDPDQYGNCRQVKAHAKFNHICE	300			
Qy	659	QTRFKN 664				
Db	301	QTRFKN 306				

RESULT 2

C96516

F16N3.15 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001


```

Query Match          3.9%; Score 141.5; DB 2; Length 1208;
Best Local Similarity 26.1%; Pred. No. 0.42;
Matches 54; Conservative 28; Mismatches 92; Indels 33; Gaps 8
159 PVPVPSPTQVCFEFPQSTSDLFPASASSTPEPAGRRSSCGSGGGSPKGS 218
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
999 PRAP---LMTGIWALPAP---VGDAVPSAGLRSDPSINWERTD-----RLKTA 104

```

RESULT 10
EDBEXD
Immediate-early protein RL2 - human herpesvirus 2 (strain HGS2)

N:Alternate names: RL2 protein
C:Species: human herpesvirus 2
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
P:Accession: U01501
P:McGeoch, N. J.; Cunningham, C.; McIntyre, G.; Dolan, A.

J. Gen. Virol. 72, 3057-3075, 1991
A:Title: Comparative sequence analysis of the long repeat regions and adjoining parts of
A:Reference number: JQ1494; MUID:92113549; PMID:1662697
A:Accession: JQ1501
A:Molecule type: DNA
A:Residues: 1-825 <MC>
A:Cross-references: GB:D10471; DBJ:D01128; NID:G221784; PIDN:BAA23427.1; PID:G2626942
C:Genetics:
A:Gene: RL2
A:Introns: 25/3; 252/1
C:Superfamily: herpesvirus immediate-early protein IB110; RING finger homology
C:Keywords: DNA binding; immediate-early protein; tandem repeat; transcription regulation
F:122-172/Domain: RING finger homology <RNG>
F:126-166/Region: zinc finger C3HC4 motif
F:589-623/Region: 5-residue repeats (A-S-S-S-S)

Query Match 3.9%; Score 139; DB 1; Length 825;
Best Local Similarity 22.7%; Pred. No. 0.38; Mismatches 127; Indels 120; Gaps 19;
Matches 82; Conservative 33;

QY 152 GAGRGADP--RVPPSPATQVCFEPPQSTSTSDLFPTASASSTPEFPAG-----RE 201
DB 6 GTSRADPGEPFRPTQGTQPA--APHAWGLNDQWLLASSDSEETEVGISDDDLHRD 63
QY 202 RSSSCSGGGGIGSPKSGSPVASDESSPASALPDRTKATSFSSLEPDQDLEPVKKVR 261
DB 64 STSEAG-----STDTMEFEAGLMDAAT-----PPARPAPRQSG 96
QY 262 GDGDLNLGQLLVAPRRNAQTVEDVRAAGKPDQWEETLTCTICQDLHDCVSLQ--P 319
DB 97 SPTPADAQSC-----GGPGVGEAEAGGGD-----VCAVTDIAIPLRCQSP 143
QY 320 CMHTFCAACYSQWMSRLCPTCRPVERICKNHLNVLVEAYLIQHPPKSRSEEDVQSM 379
DB 144 CLHPFCIPCKMTWPLENTCPLCNTPV-----AYLIVGTASGSFSTIPTIV 189
QY 380 -DARKITQDML-----OPKVR-RSFS-----DEEGSSEDLLEL 411
DB 190 NDRTRVEAEAAVRAGTAVDFITGNPRTAPRSLSLGGHTVRLSTPFPFGTDDDDLL 249
QY 412 SDVDSSESDISQPVV--VCRQCE-----YRROAQAQPHC-PAPEGSPGAPO--ALGD 459
DB 250 ADVD-----YVPPAPRAPRRGGGAGATRGTSQAATRAPAP---PGAPRSSSSGG 297
QY 460 AP 461
DB 298 AP 299

RESULT 11
T39553
probable DNA repair and recombination protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C:Accession: T39553
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, October 1998
A:Reference number: 221868
A:Accession: T39553
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-387 <LYN>
A:Cross-references: EMBL:AL031856; PIDN:CAA21300.1; GSPDB:GN00067; SPDB:SPBC1734.06
A:Experimental source: strain 972h; cosmid c1734
C:Genetics:
A:Gene: SPBC1734.06
A:Map position: 2
A:Superfamily: RING finger homology
F:25-72/Domain: RING finger homology <RRN>

Query Match 3.8%; Score 138; DB 2; Length 387;
Best Local Similarity 25.2%; Pred. No. 0.17; Mismatches 71; Indels 14; Gaps 6;
Matches 38; Conservative 28;

QY 298 MEETLTCTICQDLHDCVSLQPCWHTFCAACYSQWMSRLCPTCRPVE--RICKNHL 355
DB 23 LDSSRLCLICHEYFR-APLITSCSTFTCSFCIRYLREHPMCAPACRAPEQESRLRKNTIL 81
QY 356 NNLVEAYLIQHPPKSRSEEDVQSMARKTKTQDMLQPKVRRSFSDRGSSSEDLLELSDVD 415
DB 82 EEILLESFKVIRPTL-----FEFLKVEN-VPKVLQAPETVIAQDSASGDEEWED--DLA 132
QY 416 SESSDISQPVVCRQCPEYRRAQAQPPHCPA 446
DB 133 SNSPASIATKTSRDSKKRKE--DLVHCPA 161

RESULT 12
T16648
hypothetical protein R02E12.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C:Accession: T16648
R:Leimbach, D.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid R02E12.
A:Reference number: Z18554
A:Accession: T16648
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-639 <LEI>
A:Cross-references: EMBL:U53337; NID:G1255833; PID:G1255835; PIDN:AAA96184.1; GSPDB:GN00067
A:Experimental source: strain Bristol N2; clone R02E12
C:Genetics:
A:Gene: CESP:R02E12.4
A:Map position: X 3.8%; Score 138; DB 2; Length 639;
A:Introns: 63/3; 90/1; 136/2; 177/1; 220/2; 242/3; 294/2; 427/2; 464/3; 505/3; 580/1
C:Superfamily: RING finger homology
F:22-70/Domain: RING finger homology <RRN>

Query Match 3.8%; Score 138; DB 2; Length 639;
Best Local Similarity 28.8%; Pred. No. 0.32; Mismatches 66; Indels 22; Gaps 7;
Matches 44; Conservative 21;

QY 283 TVHEVRAAAGKPD--KME-ETLCTICQDLHDCVSLQPCWHTFCAACYSQWMSRLC 339
DB 2 TVEEQVEMNLDPESLEFEDVQCHICFQVNHPEVFTLTCKHSICACAGRWLSSCSVC 61
QY 340 PTCRCPVERICKNHLNVLVEAYLIQHPPKSRSEEDVQ-----SMDARNKITQ- 387
DB 62 FMCRTVREIHLNDDLLKKSLFLKLPHEQ-ELPEDIKWEQCCHETVFWSLQQRKRNES 120
QY 388 DML-----QPKVRRSFSDRGSSSEDLLELSDVD 416
DB 121 DAVFYGSQPVVPEHLASAE--SNDPIETKRLN 151

RESULT 13
B89009
protein T27C4.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B89009
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/ and
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B89009
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC17700.1; PID:G3165588; GSPDB:GN00023; CESP:T27C4.4
C:Genetics:
A:Gene: T27C4.4

A:Map position: 5

Query Match 3.8%; Score 137.5; DB 2; Length 973;
Best Local Similarity 22.4%; Pred. No. 0.58;
Matches 95; Conservative 55; Mismatches 190; Indels 85; Gaps 18;

QY 72 VLEDTSTSGT--VINKL-KVVKQTCTPLOTGDIYLYVRKNEPEHNVAYLYESLSEKQG 128
DB 21 VYFDDTSATDAFYQIRKIEELVKTEKGGVDARCVVYL-RRRDIPQH-----LLKIADQQR 75
QY 129 MTQESFEANKNVFHTKDTSGAGAGRGADPRVPSPSP-ATQVCFE--EPQSTSTSLF 185
DB 76 RFDNYEVKKPKENFT--TKGFIYVNGASEAEPAAPEATEASATSETKVKEDTDVA 133
QY 186 PTASASSTPSPAGRRSSSCSGGGISPKGSGSVASDEVSSPASALPDKRTAFSSSL 245
DB 134 ERKMEDEECFAPVQGRTST-----EPPS-ASDAPAAATKDKKEKEAKENAK 181
QY 245 EPDQDRLPDKVKKWEGDGLDLNGQLLVAQPRRQAQ-----TVHEDVRAAGKFDKX- 298
DB 182 NUKETRAEATKLDWGDGGLPLGVDTKLTPODLRLKQHEIFWTQSSILPAAALRGKR 241
QY 299 -----EFTLTCTIICQDLLHDCVSLQPCMHFTCA-----ACYSGWMMERSSLC 339
DB 242 VVLGGGEAQAQNYPLDDTFYHSLVYDPAQTLADKGAIRVGEKYQAVVDEWME----- 296
QY 340 PTCRCPEVERICKN-HILNNLVEAYLIQHDPKRSSEEDVQSDARNKITQDMLQPKVRSF 398
DB 297 -----PADREAKEAKYLAKEVKEBEAKLKAEEEDTENDGL--TIAEDEMPKLKNEE 349
QY 399 SDEEGSSDLLELSDVDSSESS-----DISQPYVVCRCQPEYR--QAAQ 440
DB 350 KEETKVEE---EATTTDSGEVGLVWHPHALTDRDIDQYMIVARSVGLPARAIDGAS 406
QY 441 PPHCP 445
DB 407 APXLP 411

RESULT 14

T04938
Hypothetical protein F7J7.10 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 02-Sep-2000
C:Accession: T04938
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
Submitted to the Protein Sequence Database, July 1998
A:Reference number: Z15391
A:Accession: T04938
A:Molecule type: DNA
A:Residues: 1-1331 <BEV>
A:Cross-references: EMBL:ALQ21960
A:Experimental source: cultivar Columbia; BAC clone F7J7
C:Genetics:
A:Map position: 4
A:Note: Intron positions not resolved
A:Note: F7J7.10
C:Superfamily: RING finger homology
F:377-424/Domain: RING finger homology <RRN>

Query Match 3.8%; Score 136.5; DB 2; Length 1331;
Best Local Similarity 24.6%; Pred. No. 0.98;
Matches 49; Conservative 30; Mismatches 69; Indels 51; Gaps 7;

QY 272 LLVAQPRRNaQTVHEDVRAAG-----KPDKMEET----- 301
DB 319 MLISTPERSPTIVVKNLRVCADCHLAIKLVSKVYNREIVVRDRSRKQADTSHLERMGRE 378
QY 302 LTCIICQDLLHDCVSLQPCMHFTCAACYSGWMERSSLCPTCRCPVER--ICKNHILNNLV 359
DB 379 LKCPICLSLNSAVLS-CNHVFCNACIVKSMKMDATCPVCKIPYHREIRGAPHMDSLV 437
QY 360 EAY-----LIQHPDKRSSEEDVQSDARNKITQDMLQF---KVRSPSDEEGSS 405

DB 438 SIYKNMEDASGIKLFVSGNNPSPSKKEKQVRDASVEKASDKNRQSRKGRASKRNEYGKT 497
QY 406 EDLLELSDVDSSESSDISQP 424
DB 498 KEI-----DVDAPGPIVMKP 512

RESULT 15

A85240
Hypothetical protein AT4g21070 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: A85240
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: A85240
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1495 <STO>
A:Cross-references: GB:NC_001268; NID:g7268904; PIDN:CAB79107.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g21070
A:Map position: 4

Query Match 3.8%; Score 136.5; DB 2; Length 1495;
Best Local Similarity 24.6%; Pred. No. 1.1;
Matches 49; Conservative 30; Mismatches 69; Indels 51; Gaps 7;

QY 272 LLVAQPRRNaQTVHEDVRAAG-----KPDKMEET----- 301
DB 536 MLISTPERSPTIVVKNLRVCADCHLAIKLVSKVYNREIVVRDRSRKQADTSHLERMGRE 595
QY 302 LTCIICQDLLHDCVSLQPCMHFTCAACYSGWMERSSLCPTCRCPVER--ICKNHILNNLV 359
DB 596 LKCPICLSLNSAVLS-CNHVFCNACIVKSMKMDATCPVCKIPYHREIRGAPHMDSLV 654
QY 360 EAY-----LIQHPDKRSSEEDVQSDARNKITQDMLQF---KVRSPSDEEGSS 405
DB 655 SIYKNMEDASGIKLFVSGNNPSPSKKEKQVRDASVEKASDKNRQSRKGRASKRNEYGKT 714
QY 406 EDLLELSDVDSSESSDISQP 424
DB 715 KEI-----DVDAPGPIVMKP 729

Search completed: May 7, 2004, 14:50:06
Job time : 40.4116 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:37:32 ; Search time 22.4372 seconds
(without alignment)

1540.951 Million cell updates/sec

Title: US-10-048-046-2

Perfect score: 3585

Sequence: 1 MERPEGKQSPFPQNGRLL.....VKAHAMKFNHCETRFKN 664

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	203.5	5.7	485	1 RNF8 HUMAN	O76064 homo sapien
2	197	5.5	488	1 RNF8 MOUSE	O8vc56 mus musculus
3	168.5	4.7	467	1 ICP0 VZV	P09309 varicella-z
4	145	4.0	1603	1 PSC DROME	P35820 drosophila
5	139	3.9	775	1 ICP0 HSV11	P08393 herpes simp
6	139	3.9	825	1 ICP0 HSV2H	P28284 herpes simp
7	138	3.8	387	1 RH18 SCHPO	O74747 schizosacch
8	134.5	3.8	551	1 RN27 HUMAN	O9bzt9 homo sapien
9	134	3.7	624	1 RN12 HUMAN	O9nvw2 homo sapien
10	132.5	3.7	1300	1 SAL3 HUMAN	O9bxa9 homo sapien
11	132.5	3.7	2414	1 P300 HUMAN	Q09472 homo sapien
12	131.5	3.7	600	1 RN12 MOUSE	Q9wtv7 mus musculus
13	131	3.7	1849	1 IGA4 HAIN	P45386 haemophilus
14	130.5	3.6	551	1 RN27 MOUSE	O99pj2 mus musculus
15	129	3.6	1341	1 ACIN2 HUMAN	O9ukv3 homo sapien
16	128.5	3.6	933	1 PRGR HUMAN	P06401 homo sapien
17	128	3.6	407	1 RPF2 HUMAN	O80858 homo sapien
18	128	3.6	1338	1 ACIN2 MOUSE	Q9jix8 mus musculus
19	128	3.6	2003	1 NTC4 HUMAN	O99466 homo sapien
20	127	3.5	2715	1 MLL4 HUMAN	O9umt6 homo sapien
21	126.5	3.5	676	1 ICP0 HSVBJ	P29128 bovine herp
22	126	3.5	744	1 TRM3 HUMAN	O75382 homo sapien
23	125.5	3.5	423	1 MKR2 SERQU	O9dd48 seriola qui
24	125.5	3.5	989	1 SRAA FLAUF	P32823 plasmodium
25	125	3.5	630	1 TRM3 HUMAN	Q4258 homo sapien
26	125	3.5	744	1 TRM3 RAT	O70277 rattus norv
27	125	3.5	777	1 BAR1 HUMAN	Q99728 homo sapien
28	125	3.5	794	1 ICA1 BOVIN	O60821 mus musculus
29	124.5	3.5	705	1 ICA1 BOVIN	P20811 bos taurus
30	124.5	3.5	1127	1 TRIG HUMAN	O9upn9 homo sapien
31	124.5	3.5	3866	1 HRX MOUSE	P55200 mus musculus
32	124	3.5	676	1 ICP0 HSVBK	P29536 bovine herp
33	124	3.5	1712	1 L7B1 RAT	Q00918 rattus norv

34	124	3.5	2842	1 APC RAT	P70478 rattus norv
35	123.5	3.4	892	1 ATX7 HUMAN	O15265 homo sapien
36	123.5	3.4	1233	1 M4K4 MOUSE	P97820 mus musculus
37	123	3.4	605	1 PJAI MOUSE	O55176 mus musculus
38	123	3.4	898	1 CIZ1 HUMAN	O9ulv3 homo sapien
39	123	3.4	1007	1 SAL2 HUMAN	Q9y457 homo sapien
40	122	3.4	433	1 TIG CHLMU	Q9p119 chlamydia m
41	122	3.4	501	1 UVS2 NEUCR	P33288 neurospora
42	122	3.4	507	1 MKR3 HUMAN	Q13064 homo sapien
43	122	3.4	605	1 CORO CAREL	Q21624 caenorhabdi
44	122	3.4	1004	1 SAL2 MOUSE	O9qx96 mus musculus
45	122	3.4	1073	1 PVDA FLAKN	P22545 plasmodium

ALIGNMENTS

RESULT 1	RNF8_HUMAN	STANDARD;	PRT;	485 AA.
ID	RNF8_HUMAN			
AC	O76064;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	RING finger protein 8.			
GN	RNF8 OR KIAA0646.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	Seki N., Yamauchi M., Saito T.;			
RT	"Isolation and chromosomal assignment of the gene for a novel zinc			
RT	finger protein.";			
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	MEDLINE=98403880; PubMed=9734811;			
RA	Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,			
RA	Kotani H., Nomura N., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. X.			
RT	The complete sequences of 100 new cDNA clones from brain which can			
RT	code for large proteins in vitro.";			
RL	DNA Res. 5:169-176 (1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	Tracey A.;			
RA	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Muscle;			
RA	MEDLINE=22388257; PubMed=12477932;			
RA	Stausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Schestz T.E.,			
RA	Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Faney J., Heiton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakeley R.W., Touchman J.W., Green B.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			

RT	human and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC	-I- SIMILARITY: Contains 1 PHA domain.
CC	-II- SIMILARITY: Contains 1 RING-type zinc finger.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to licensee@isb-sib.ch).
CC	-----
DR	EMBL; AB012770; BAA33557.1; -
DR	ENBL; AB014546; BAA31621.1; -
DR	ENBL; AL096712; CAB75689.1; -
DR	ENBL; BC007517; AAK07517.1; -
DR	Genew; HGNC:10071; RNFB.
DR	InterPro; IPR000253; FHA.
DR	InterPro; IPR008984; SMAD_FHA.
DR	InterPro; IPR001841; Znf_Ring.
DR	Pfam; PF00498; FHA; 1.
DR	Pfam; PF00097; zf-C3HC4; 1.
DR	SMART; SM00240; FHA; 1.
DR	SMART; SM00184; RING; 1.
DR	PROSITE; PSS0006; FHA_DOMAIN; 1.
DR	PROSITE; PS00518; ZF_RING_1; 1.
DR	PROSITE; PS50089; ZF_RING_2; 1.
KW	Zinc-finger.
FT	DOMAIN 38 92 FHA.
FT	DOMAIN 276 345 GLN-RICH.
FT	DOMAIN 403 441 RING-TYPE.
ZN	FING
SS	SEQUENCE 485 AA; 55517 MW; 54650B2FFC9948B1 CRC64;
SQ	-----
Query Match	5.7%; Score 203.5; DB 1; Length 485;
Best Local Similarity	22.5%; Pred. No. 2.2e-05;
Matches 102; Conservative	60; Mismatches 170; Indels 121; Gaps 17;
QY	19 LLRLGAEGEHPVLLRKR-EWTIGRRRGCDLSPFSN--KLVSGDHCRIVVDKSGGVTL 74
DB	20 LRRVGMSAG--WLLEPGCEVTGVRGFVTYQLVSKICPLMISRNHC-VLKQNPFGOWTI 76
QY	75 EDT-STSTGTWINKVKVKKOTCPLOTGDVIYL-----VYRKNEPEHNVAIYESL 123
DB	77 MDNLSLNGVMNRARLEPLRVSIHQDYIQLGVPLENKENARYEVEVTEDWETIYPCL 136
QY	124 SEKGQMTQESFEANKNVFHCTKDTCAGAGRGADRPVPSSPATQVCPEPOP----- 177
DB	137 SPK---NDQMIEKNKE-LRTYRKFSDELAPGAGPSNLKSKINKVCSGGDPVKSQCK 192
QY	178 ----STSTDLPPTASA---SSTEPPSA-----GRERSSSCSGGGGI----- 213
DB	193 GEVASTPDSNLDPKLTALFPSTKTGATIPGPFKVTEVHHHEKAASNSASQSRLQMPKVT 252
QY	214 -----SPKSGSPSVAS-----DEVSSFSALPDPRKTASFSS 244
DB	253 MSRIILRIKIOMEKHEAVMNVKQTQGNSEKVVQMEQLDLOSQLCBAEQAOQAARVEQ 312
QY	245 LEPOQDEDLPEPVKKMRGDGLDNLGQL-----LVQPFRN 280
DB	313 LEKTFEEQECHLOGLETAQGEKDLKQLLAQALQEHWMELBNRSKKDFRAIIQAKNEL 372
QY	281 AQTVHDVRAARGP-----DKWEETLTICIIDLHDCVSLQPCWHHFCAACYSGWM 333
DB	373 EOTKEEKMQAQKEEVLSHWMDVLNDELQCIIISEYFIEAVTLN-CAHSFCSCINEMW 431
QY	334 ERSSLCEPTCRCFVER-----ICKNIHLNNL 358
DB	432 KRKEICTPCRDKDIKSTYSVLIDNCINKWNANL 464

RESULT 2
RNF8 MOUSE

ID	RFN8_MOUSE	STANDARD;	PRT;	488 AA.
DT	QBVC56;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	RING finger protein 8.			
GN	RFN8			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM M.A.			
RC	TISSUE=kidney;			
XX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Uedtin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S.E., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Pahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S., Krzywinski M.I., Skalska U., Smallos D.E.,			
RA	Schmerer A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences";			
RL	Proc. Natl Acad. Sci U.S.A. 99:16899-16903(2002).			
CC	-1- SIMILARITY: Contains 1 FHA domain.			
CC	-1- SIMILARITY: Contains 1 RING-type zinc finger.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; BC021778; AAH21778.1; -.			
DR	MCD; MGI:1929069; Rnf8.			
DR	InterPro; IPR000253; FHA.			
DR	InterPro; IPR008984; SMAD_FHA.			
DR	InterPro; IPR001841; Znf_Ring.			
DR	Pfam; PF00498; FHA; 1.			
DR	Pfam; PF00097; zf-C3HC4; 1.			
DR	SMART; SM00240; FHA; 1.			
DR	SMART; SM00184; RING; 1.			
DR	PROSITE; PS00066; FHA DOMAIN; 1.			
DR	PROSITE; PS00518; ZF_RING_1; 1.			
DR	PROSITE; PS00089; ZF_RING_2; 1.			
KW	Zinc-finger.			
FT	DOMAIN 38 92 FHA.			
FT	DOMAIN 279 345 GLN-RICH.			
FT	ZN FING 406 444 RING-TYPE.			
FT	SEQUENCE 488 AA; 55516 MW; 428242204BC44A1 CRC64;			
QY	Query Match			
DB	Best Local Similarity 21.0%; Score 197; DB 1; Length 488;			
QY	Matches 98; Conservative 55; Mismatches 173; Indels 140; Gaps 17;			
QY	27 GEPHVL-----RKREWTIGRRGCD-----LSFSPSKLWS 57			
DB	2 GEPDPLVSGQLAARRSWCL--RLGLMCEWQLQLEAGTEVTIGRLSVTYQLISKVCPLMIS 60			
QY	58 GDHCRIIVVDEKSGQVLTEDT-STSGTVINKLKVKKQTCPLQTDGVIVLYVRKNEEHNV 116			

Db 61 RSHC-VLKONPEQWIMDNKSLNGVWLNRLAPLQGYCIRKGDHIOQLGV-PLESRETA 118
QY 117 AYLYESLSE-----KQGWTOBSFEA-NKENV-----141
Db 119 EYEVIEEDWESLAPCLAPKNDQRMKEKHGSRTKKFSFSGLENI-PARGSSDLRCPLAN 178
QY 142 -----FHGTQYTSAGAG-----RGADPRVPPSPAT-----Q 169
Db 179 VASKPIPEXKLGHGDASSQSLGCLPGLTSLKASERAAAGPHACSLPKVLELSLCPKKQK 238
QY 170 VCPEEPQSTSDLPFTASASTESPAGRRSSSCGGGGISPKGSGSPVASDE--- 226
Db 239 AC--RPSASQNSUELFKVTWMSRLKLTQMOEQIAV-LNVKQTRKGSKKIVRNEKEL 295
QY 227 --VSSFASALPDRKTASFSSLEFQDQDLEBPVKKMKRGDGLDLNGLQLVVAOPRRA--- 281
Db 296 RNLQSLYASQAQQAARVEQLTQFEAHYLOGLKEQECGLKQQLVQALQEQHQLME 355
QY 282 -----QTVHEDVRAAAGKP-----DKMEETLTCTICQDQLLHD 313
Db 356 ELNCSKDFEKKIQAOKNKELEQTKKEKDKVQAQKEVLSHMNDLLENELQCIICSEYFIE 415
QY 314 CVSLQPCWHTFCAACYSQWMMERSLCTCRCPVERICKHILNNLV 359
Db 416 AVTLN-CAHSFCFCFNEWMKRVCEPICRDKIESRTNSILVLDNCI 460
RESULT 3
ICPO_VZVD STANDARD; PRT; 467 AA.
AC P09309;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-acting transcriptional protein ICPO.
GN 61.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93059681; PubMed=1366099;
RA Moriuchi H., Moriuchi M., Smith H.A., Straus S.E., Cohen J.I.;
RA "Varicella-zoster virus open reading frame 61 protein is functionally
RT homologous to herpes simplex virus type 1 ICPO.";
RL J. Virol. 66:7303-7308 (1992).
CC -!- FUNCTION: REPRESSES THE EXPRESSION OF VIRAL IE, EARLY, AND LATE
CC GENE PROMOTERS.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES ICPO PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04370; CAA27944.1; --
CC PIR; I27215; WZBE61.
CC HSPSP; P28990; ICHC.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00097; zf-C3HC4; 1.
CC SMART; SMO0184; RING; 1.

DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Transcription regulation; Trans-acting factor; Repressor; Zinc-finger;
KW DNA-binding.
FT ZN-FING 19 58 RING-TYPE.
SQ SEQUENCE 467 AA; 50916 MW; 25EFA6977EA6994C CRC64;
Query Match 4.7%; Score 168.5; DB 1; Length 467;
Best Local Similarity 26.3%; Pred. No. 0.0028;
Matches 64; Conservative 20; Mismatches 106; Indels 53; Gaps 10;
QY 287 DVRAAGKPDKNMEETLTCTICQDQLLHDCVSLQPCWHTFCAACYSQWMMERSLCTCRCPV 346
Db 2 DTLAGSGSTSDASDNTCTICWSTVSDLGKTMPCLDHDFCVICIRAWTSTSVQCPLCRCPV 61
QY 347 ERICKHILNNLVAYLIQHPDKSRSEEDVQSDARKKITQDMLQPKVRRFSDEEGSSE 406
Db 62 QSILFKIVSDTSYKEYEV-HP-----SDDGFS-----EPSFSDIDILPGDVI 104
QY 407 DLLESLDVDSSESDISQPVVYVCRQCPQEVRRQAAOPPHCPAPEGEPGAPQALCDAPSTSVS 466
Db 105 DLLPSPGSPRES-IQPP-----TSRSRPIQSP-----NPGPLQSSARBP-TAES 149
QY 467 LTTAVQDYVCP--LQSHALCTCC-----FQMPDPRRA-----EREQDRV 505
Db 150 PSDSQDSIQPPTRDSSPGVTKCTGASTFLRKVFKDQPAVRSATPVVYVGSIESAQQPR 209
QY 506 APQ 508
Db 210 GQQ 212
RESULT 4
PSC_DROME
ID_PSC_DROME STANDARD; PRT; 1603 AA.
AC P35820;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Posterior sex combs protein.
GN PSC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92018190; PubMed=1833647;
RA Brunk B.P., Martin E.C., Sharp E., Adler P.N.;
RA "Drosophila genes Posterior Sex Combs and Suppressor two of zeste
RT encode proteins with homology to the murine bmi-1 oncogene.";
RL Nature 353:351-353 (1991).
CC -!- FUNCTION: The Polycomb group (PC-G) genes are needed to maintain
CC expression patterns of the homeotic selector genes of the
CC antennapedia (Antp-C) and bithorax (Bx-C) complexes, and hence for
CC the maintenance of segmental determination.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X59275; CAA41965.1; --
CC PIR; S17983; S17983.
CC FlyBase; FBgn0005624; Psc.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00097; zf-C3HC4; 1.


```

J. Gen. Virol. 69:1531-1574 (1988).
[2] SEQUENCE FROM N.A.
RX MEDLINE=87059760; PubMed=3023529;
RA Perry L.J., Rixon F.J., Everett R.D., Frame M.C., McGeoch D.J.;
RT "Characterization of the IE110 gene of herpes simplex virus type 1.";
RL J. Gen. Virol. 67:2365-2380 (1986).
[3] SEQUENCE FROM N.A.
RX MEDLINE=89036163; PubMed=2846760;
RA Perry L.J., McGeoch D.J.;
RT "The DNA sequences of the long repeat region and adjoining parts of
the long unique region in the genome of herpes simplex virus type
1.";
RL J. Gen. Virol. 69:2831-2846 (1988).
[4] FUNCTION: ICPO IS A GENERAL TRANSACTIVATOR OF ALL THREE CLASSES OF
CC HSV GENES WHICH FUNCTIONS SYNERGISTICALLY WITH ICPO4. OFTEN TERMED
CC AS A PROMISCUOUS TRANSACTIVATOR. MAY PLAY A ROLE IN THE BALANCE
CC BETWEEN THE LATENT AND LYtic STATES, REACTIVATING LATENT HSV. ICPO
CC MIGRATES TO THE ND10 NUCLEAR BODIES IN THE EARLY TIMES OF
CC INFECTION AND DISRUPTS THEM. BINDS TO HAUSP AND MAY MODIFY ITS
CC DEURBUTINATING SUBSTRATE SPECIFICITY OR ACTIVITY ON VIRAL OR
CC CELLULAR TARGETS, LEADING TO AN INCREASED OR DECREASED STABILITY
CC OF THESE PROTEINS.
CC [5] SIMILARITY: Contains 1 RING-type zinc finger.
CC [6] SIMILARITY: TO OTHER HERPESVIRUSES ICPO PROTEIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation on
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC EMBL; X14112; CAA32336.1; -
DR EMBL; X14112; CAA32293.1; -
DR EMBL; X04614; CAA28285.1; -
DR PIR; A29152; EDBE11.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Transcription regulation; Trans-acting factor; Activator; Zinc-finger;
FT DNA-binding; Early protein.
FT ZN FING 116 157 RING-TYPE.
FT DOMAIN 233 243 POLY-ASP.
FT DOMAIN 305 308 POLY-GLY.
FT DOMAIN 558 568 POLY-SER.
SQ SEQUENCE 775 AA; 78456 MW; DF38ALC539DAB15C CRC64;
Query Match 3.9%; Score 139; DB 1; Length 775;
Best Local Similarity 22.6%; Pred. No. 0.33;
Matches 80; Conservative
Qy 156 GADPRVPPSPATQVCFEEPQST----SISLFFPTASASTPSPAGRSSSCSGGG 211
Db 6 GASTRRDEGRQ-----REFAPDVVFPQDRDL-PPSSDSEAEVEVGR----- 48
Qy 212 GISPKGSGPSVADSVSSFSALPDRKTASFSLSLEPQDLELPVKKKRGDGLDLNGQ 271
Db 49 -----GDADHDD-----DSASEADSTDTELFETGLGPGQVD-GGA 84
Qy 272 LLVAQPRNQAQTVHEDYRAAGKPKD---MEEFLTCIIQDILL--HDCVSLQPCMHFTC 325
Db 85 VSGSGSPR-----EEDPGSCGGAPPREDGSGDEGDVCAVCTDETAFLHRCDTFFCMHRC 139
Qy 326 AACYSYGMWSSSLCPTCRCPVERICKNHLNLLNVEAYLIQHPDKR--SEEDVOSMDARN 383
Db 140 IPCMKTWQLRNTCFLCNALVLVLIVGVTPSGSPSTPIVNDPQTRMEAEAVRAGTAVD 199
Qy 384 KI-----TQDMLQPKVRFSFSDSEGSDELLEISDVDSSESSDISQPV 426

```

Db 200 FIWGNORFAPRYLTGHTVRLSPHPTPTDDE-----DDLDLDADYVPP--- 247
QY 427 VCQCCEY--RRQAOPP-----HCPAPEGGAQA--LGDAPSTSVSLTT 469
Db 248 APRTPRPRRGAAPPVTTGGSHAAPQAAARTAPPAPIGPHGSSNTNTTT 301
RESULT 6
ID_ICPO_HSV2H STANDARD; PRT; 825 AA.
AC P28284;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-acting transcriptional protein ICPO (VNM118 protein).
GN RL2.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.,
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2."
RN J. Gen. Virol. 72:3057-3075(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Dolan A.,
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ICPO PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D10471; BAA23427.1; -
DR EMBL; Z86099; CAB06760.1; -
DR PIR; JQ1501; EDBEXD.
DR HSSP; P28990; ICHC.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Transcription regulation; Trans-acting factor; Activator; Zinc-finger;
KW DNA-binding.
FT DOMAIN 120 123 POLY-GLY.
FT ZN_FING 126 167 RING-TYPE.
FT POLY-GLY 266 271 POLY-GLY.
FT DOMAIN 292 295 POLY-SER.
FT DOMAIN 342 345 POLY-ALA.
FT DOMAIN 386 389 POLY-SER.
FT DOMAIN 395 400 POLY-GLY.
FT DOMAIN 425 428 POLY-ALA.
FT DOMAIN 590 627 POLY-SER.
SQ SEQUENCE 825 AA; 81986 MW; 5CEB1585553A274 CRC64;
Query Match 3.9%; Score 139; DB 1; Length 825;
Best Local Similarity 22.7%; Pred.No.0.36;
Matches 82; Conservative 33; Mismatches 127; Indels 120; Gaps 19;
QY 152 GAGRGADP---RVPPSSPAIVCFEPPQSTSTDLFTPTASSTEPSAG-----RE 201
Db 6 GTSSRADPGPPPPPTGTCTQPA---APHAWGLNDQMQLASSDSSEETEVGISDLDLHRD 63

QY 202 RSSSCGSGGGISPKSGSPSVASDEVSSFASALPDRKTAESPSSLEPOQEDLEPVKKMR 261
Db 64 STSEAG-----STDTEMFEAGLMDAAT-----PPRPPAERQG 96
QY 262 GDGDLDLNGQLLVAQPRRNAQTVHEDVRAAAGKPKMEETLTCTIICODLLHDCVSLQ--P 319
Db 97 SPTPADAGSC-----GGPVGEEAEAGGGD-----VCAVCTDEIAPPLRCQSF 143
QY 320 CMHTFCAACYSGNMRSSLCPTCRCPVERICKNHLNVLVEAYLIQHPDKSRSEDEVQSM 379
Db 144 CLHPFCIPCKMTWPLRNTCLNTFV-----AYLIVGVTAGSGSTIPV 189
QY 380 -DARNKITQDML-----QPKVR-RSFS-----DEGSGSDDLLEL 411
Db 190 NDRTRVRAEAAVRAGTAVDFWNTGNPTAPRSLSLGHTVRLSPHPTPTDDEDDL 249
QY 412 SDVDSSESDISQPVV--VCRCPE-----YRQAQAPPHC-PAPEGEPAQ--ALGD 459
Db 250 ADVD-----YVPPAPRAPRRGGGAGATGTSQPAATRPAP---PGAPRSSSGG 297
QY 460 AP 461
Db 298 AP 299
RESULT 7
RH18_SCHPO
ID_RH18_SCHPO STANDARD; PRT; 387 AA.
AC O74747;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Postreplication repair protein rhp18 (RAD18 homolog).
GN RHP18 OR SPBC1734.06
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=2155918; PubMed=11702950;
RA Kitamura K., Katayama S., Dhut S., Sato M., Watanabe Y., Yamamoto M.,
RA Toda T.,
RT "Phosphorylation of Mei2 and Ste11 by Pat1 kinase inhibits sexual
RT differentiation via ubiquitin proteolysis and 14-3-3 protein in
RT fission yeast."
RL Dev. Cell 1:389-399(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayles J., Baker S., Chillingworth T., Churcher C.M.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Fraser A.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellon J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B., Auer S.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motter S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

DR PROSITE; PS50089; ZF_RING 2; 1.
 KW Nuclear protein; Coiled coil; Zinc-finger; Repeat.
 FT ZN FING 15 56 RING-TYPE
 FT ZN FING 92 132 B BOX-TYPE 1
 FT ZN FING 140 182 B BOX-TYPE 2
 FT DOMAIN 181 249 COILED COIL (POTENTIAL).
 FT CONFLICT 174 174 H -> R (IN REF. 1).
 SQ SEQUENCE 551 AA; 61489 MW; 1FEF89029BE9BAC CRC64;
 Query Match 3.8%; Score 134.5; DB 1; Length 551;
 Best Local Similarity 23.8%; Pred. No. 0.41; Mismatches 87; Gaps 11;
 Matches 58; Conservative 20;
 QY 299 EETLTICQDLHDCVSLQPCMHTEFAACY-SGMWRSLL--CPTCRPVER---ICKN 352
 Db 10 EEEELICPLHVFVEPVL-PCKHNEFCRGIGBAAKDSGLVRCPECNAYNQKPLEKN 68
 QY 353 HILNNLVEAYLQHPDKS-----RSEEDVQSDAENKIKTDWL 390
 Db 69 LKLTNIVEKFNALHVEKPAALHVCVRGPPPLPAQKVCIRCEAPC---CQSHVQTHLQ 124
 QY 391 QPKVRSFSDSEGSDDLLELSVDSESSDISOPYVVCRCPEYRQAQPPHCPAPEGE 450
 Db 125 QPSTARGLH-----LYEADDVRAWS-----CFQH--NAVRLVCEAEQ-- 160
 QY 451 PGAPQALGAPSTSVSLTTAVQDVYVCPLOGSHALCTCC-----FQMPDPRRAE 498
 Db 161 -----VAVQCYCCYSGAHQHSVCDVEIRRNRIKMLKQDRLLEE 202
 QY 499 REQD 502
 Db 203 REQD 206
 RESULT 9
 ID RN12 HUMAN STANDARD; PRT; 624 AA.
 AC Q9V598;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE RING finger protein 12 (LIM domain interacting RING finger protein)
 DE (RING finger LIM domain-binding protein) (R-LIM) (NY-REN-43 antigen).
 GN RNF12 OR RLIM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=20459411; PubMed=11013082;
 RX Ostendorff H.P., Bossenz M., Mincheva A., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A., Richter P., Bach I.,
 RT "Functional characterization of the gene encoding RLIM, the
 RT corepressor of LIM homeodomain transcription factors."
 RL Genomics 69:120-130(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Renal cell carcinoma;
 RX MEDLINE=99438124; PubMed=10508479;
 RA Scanlan M.J., Gordon J.D., Williamson B., Stockert E., Bander N.H.,
 RA Jongeneel C.V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,
 RA Old L.J.;
 RT "Antigens recognized by autologous antibody in patients with renal-
 RT cell carcinoma."
 RL Int. J. Cancer 83:456-464(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Magateuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,

RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RL "NEDO human cDNA sequencing project."
 CC Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Acts as a negative co-regulator for LIM homeodomain
 CC transcription factors. Via the recruitment of the SIN3a/histone
 CC deacetylase corepressor complex.
 CC -!- SUBUNIT: Associates with LIM/homeobox factors.
 CC -!- TISSUE SPECIFICITY: Expressed in many tissues.
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts
 CC in position 134 and 142.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AJ271670; CAC14228.1; .
 DR EMBL; AF155109; AAD42875.1; ALT_FRAME.
 DR EMBL; AK001334; BAA91632.1; .
 DR Genew: HGNC:13429; RNF12.
 DR MIM; 300379; .
 DR GO; GO:0017053; C:transcriptional repressor complex; NAS.
 DR GO; GO:0003714; F:transcription co-repressor activity; NAS.
 DR GO; GO:0016481; P:negative regulation of transcription; NAS.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
 DR PROSITE; PS0089; ZF_RING_2; 1.
 KW Transcription regulation; Zinc-finger.
 FT DOMAIN 422 506
 FT ZN FING 570 611 RING-TYPE.
 FT DOMAIN 453 481 POLY-SER.
 FT DOMAIN 500 506 POLY-SER.
 FT CONFLICT 126 126 S -> C (IN REF. 1).
 FT CONFLICT 134 134 N -> D (IN REF. 3).
 FT CONFLICT 144 145 YS -> NR (IN REF. 3).
 FT CONFLICT 418 418 Y -> H (IN REF. 3).
 SQ SEQUENCE 624 AA; 68527 MW; DE3ADE09ACACBCEFC8 CRC64;
 Query Match 3.7%; Score 134; DB 1; Length 624;
 Best Local Similarity 22.3%; Pred. No. 0.52;
 Matches 55; Conservative 34; Mismatches 106; Indels 52; Gaps 9;
 QY 100 GDVYLYVRKNEPEHNVAIYSELSEKQMTQESFEANKENVFHTKDTSGAGRGADP 159
 Db 414 GELSYFYMYSDSDSEPT-----GSVSNENMERAES-----RSGRGGSGGGSSSSSS 459
 QY 160 RVPSSPATQVCFEPOPSTSTDLFTASASSTEPSAGRERSSSCGGGGISPKSG 219
 Db 460 SSSSSSSSSSSSSSSSSSSSGG---ESSETSDLPESNEGSSSSSGSGARREGHRA 516
 QY 220 PSVASDEVSSPASLPDRKTASFSLPQOEDELPVKKKRGDGLDLNQLLVAQPR- 278
 Db 517 P-VTFDE---SGSLPFLSLAQFFLLNEDDDD-----QPRG 547
 QY 279 RNAQTVHEDVRAAGKDPKMEETLTCTICQDLLHDCVSLQ--PCWHTFCAACYSGMWS 336
 Db 548 LTKSQIDNLAVRSFGENDALK---TCSVCITTEYTEGNKRLKLPCHSEYHVHCIDRWLSEN 604
 QY 337 SLCPCTCR 343
 Db 605 STCPICR 611
 RESULT 10
 SAL3 HUMAN
 ID SAL3 HUMAN STANDARD; PRT; 1300 AA.

Q9BXA9; Q9UGH1;
 28-FEB-2003 (Rel. 41, Created)
 28-FEB-2003 (Rel. 41, Last sequence update)
 10-OCT-2003 (Rel. 42, Last annotation update)
 Sal-like protein 3 (zinc finger protein SALL3) (hsSALL3).
 SALL3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORM 3).
 RA Gooding R., Angelicheva D., Blechschmidt K., Swoboda K., Molnar M.,
 RA Tournev I., Kalaydjieva L.;
 RT "Exclusion of HSALL3 and refinement of the region for the CCFDN
 gene.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE OF 29-1300 FROM N.A., AND ALTERNATIVE SPLICING.
 RP MEDLINE=20079154; PubMed=10610715;
 RA Kohlbase J., Hausmann S., Stojmenovic G., Dixkens C., Bink K.,
 RA Schulz-Schaeffer W., Altmann M., Engel W.;
 RT "SALL3, a new member of the human spalt-like gene family, maps to
 RT 18q23.";
 RL Genomics 62:216-222(1999).
 [3]
 RN VARIANT LEU-593, AND REVISIONS TO 787; 797-802; 808; 1138 AND 1141.
 RP Kohlbase J.;
 RL Submitted (JUL-2002) to Swiss-Prot.
 CC -!- FUNCTION: Probable transcription factor.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=3;
 CC IsoId=Q9BXA9-1; Sequence=Displayed;
 CC Name=1;
 CC IsoId=Q9BXA9-2; Sequence=VSP_006833;
 CC Note=Lacks two zinc finger domains. Major isoform with isoform
 2;
 CC Name=2;
 CC IsoId=Q9BXA9-3; Sequence=VSP_006832; VSP_006833;
 CC Note=Lacks two zinc finger domains. Major isoform with isoform
 1;
 CC Name=4;
 CC IsoId=Q9BXA9-4; Sequence=VSP_006832;
 CC -!- TISSUE SPECIFICITY: Widely expressed in adult with highest levels
 in heart. Expressed in fetal brain (in neurons of hippocampus,
 cortex, mediodorsal and ventrolateral thalamic nuclei, putamen,
 cerebellum and brainstem).
 CC -!- DEVELOPMENTAL STAGE: In fetal brain of the 24th gestational week.
 CC -!- SIMILARITY: BELONGS TO THE SALL FAMILY OF C2H2-TYPE ZINC-FINGER
 PROTEINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; AF347021; AAK18311.1; .
 CC EMBL; AJ007421; CAB65124.1; ALT_SEQ.
 CC HSPSP; P07248; IARE.
 CC Gnew; HGNC:10527; SALL3.
 CC MIM; 605079; .
 CC InterPro; IPR007087; Znf_C2H2.
 CC Pfam; PF00096; zf-C2H2; 10.
 CC SMART; SM00355; Znf_C2H2; 10.
 CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
 CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 8.
 KW Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger;
 Metal-binding; Repeat; Alternative splicing; Polymorphism.

FT	ZN_FING	420	442	C2H2-TYPE.
FT	ZN_FING	448	470	C2H2-TYPE.
FT	ZN_FING	679	701	C2H2-TYPE.
FT	ZN_FING	707	729	C2H2-TYPE.
FT	ZN_FING	739	761	C2H2-TYPE.
FT	ZN_FING	977	999	C2H2-TYPE.
FT	ZN_FING	1005	1027	C2H2-TYPE.
FT	ZN_FING	1113	1135	C2H2-TYPE.
FT	ZN_FING	1141	1163	C2H2-TYPE.
FT	DOMAIN	198	201	POLY-ALA.
FT	DOMAIN	213	216	POLY-GLN.
FT	DOMAIN	897	902	POLY-SER.
FT	VARSPLIC	1	133	Missing (in isoform 2 and isoform 4). /FtId=VSP_006832.
FT	VARSPLIC	973	1044	Missing (in isoform 1 and isoform 2). /FtId=VSP_006833.
FT	VARIANT	593	593	V->L. /FtId=VAR_014132.
FT	CONFLICT	235	235	R->C (IN REF. 2). /FtId=VAR_014132.
FT	SEQUENCE	1300	AA; 135371	MM; 82CF3BDCB6D59150
SC	SEQUENCE	1300	AA; 135371	MM; 82CF3BDCB6D59150

Query Match 3.7%; Score 132.5; DB 1; Length 1300;
 Best Local Similarity 21.5%; Pred. No. 1.6;
 Matches 134; Conservative 61; Mismatches 220; Indels 209; Gaps 34;

QY	127	QGMTQESFEANKENVFR	-----GTKDTSGAGAGRGADPRVPFSSSPATQVCFE	173
Db	460	KGNLKVFQKHEKYPHIQMNPVPEYLDNVETCSGIPYGM	-----LPPEKPVTTWLD	515
QY	174	EP-QPSTSTSD--DLFT	-----ASASSTPSPAGRERSSCGSGGGI	213
Db	516	KVLTPTVTSVGLQPLPTVPGAHYADSPATPASRSPQSP	-----SEACSLSGLN	571
QY	214	-----SPKG--SGPSVASDE	-----VSFASALPDRKTASFSS	245
Db	572	HVESGVSAATRESQSLGPPVTKAEVPVSLPCTNARAGDAPVGAQAASAPTSVDGAPTSL	-----	631
QY	246	EPQDQEDLEPVKKMGDGLDLGALLVAQPRNAQTVEDVRAAGKPDKMEETLTC	-----	304
Db	632	GSPGLPAVSEQFAQFPFGGLD--SMQTSKLOQLV-ENI	-----DKMTDPNQC	681
QY	305	IICODLLHDCVSLQPCM-----HT-----FCAACYSGWMERSL	-----CPTCR-	343
Db	682	VIC-----HRVLSQCSALKMHYRTHGERPFCKICGRAFTTKGNLTKTHFGVHRAKPLRV	-----	737
QY	344	---CPVERICKHNLNLVAYLIQH-----PKSRSEEDVQSMADARKITQDMLQ	-----	391
Db	738	QHSCP---ICQKFTNAV--LQCHIRMGMGGQIPNTPLPEGFQDAM--SELAYDDKN	-----	789
QY	392	PKVRSFSD--EGSSDLELSDVDSESDISOPYV-VCRQCPYERQAAQPP	-----	442
Db	790	AETLSSYYDDMDENSMEDDAELKDA---ATDPAKFLLSYAGSCP	-----PSPPSVISSI	840
QY	443	-----HCPAPGEPGAPGALGDAPSTSVSLTAVQDYVCPQGGSHAL	-----	484
Db	841	AALENQMKMIDSVMSCOOLTGLKSVENGSGESDRLSNDSSAVGDLSESRAGSPALSESS	-----	900
QY	485	CTCCQFQMPDR-----RAEREQDPVPAQCAVCLQ	-----	515
Db	901	SSQALSAPNSGESFRSKSPCLGAPEEPQIPPLKTERPDSFAAAGSGGAPGAGIKKEA	-----	960
QY	516	PFCHLYWG-----CSTRGCVGCLAPP-CELNLDGKCDGLVNNNSYED	-----ILKNY	563
Db	961	PFSLFLSRERKCPSTVCGVCGKPFACKSAL-----EIHVRSHTKERPFVCAICRRG	-----	1013
QY	564	LATRLTWKMLTESLVALQGVF	587	
Db	1014	CSTMGNLKHLLTHRLKELPSQLF	1037	

RESULT 11
 P300_HUMAN
 ID P300_HUMAN STANDARD; PRT; 2414 AA.

AC Q09472;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE EIA-associated protein p300 (EC 2.3.1.48).
GN EP300 OR P300.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95011587; PubMed=7523245;
RA Eckner R., Ewen M.E., Newsome D., Gerdes M., Decaprio J.A.,
RA Lawrence J.B., Livingston D.M.;
RT "Molecular cloning and functional analysis of the adenovirus E1A-
RT associated 300-kD protein (p300) reveals a protein with properties of
RT a transcriptional adaptor.",
RL Genes Dev. 8:869-884(1994).
RN [2]
RN ENZYMMATIC ACTIVITY.
RX MEDLINE=97100994; PubMed=8945521;
RX Ogrzyzko V.V., Schiltz R.L., Russanova V., Howard B.H., Nakatani Y.;
RT "The transcriptional coactivators p300 and CBP are histone
RT acetyltransferases",
RL Cell 87:953-959(1996).
RN [3]
RN INTERACTION WITH PCAF.
RX MEDLINE=96300317; PubMed=8684459;
RX Yang X.-J., Ogrzyzko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.,
RA "A p300/CBP-associated factor that competes with the adenoviral
RT oncoprotein E1A",
RL Nature 382:319-324(1996).
RN [4]
RN INTERACTION WITH HIF1A AND CREBBP.
RX MEDLINE=97075102; PubMed=8917528;
RX Arany Z., Huang L.E., Eckner R., Bhattacharya S., Jiang C.,
RA Goldberg M.A., Bunn H.F., Livingston D.M.;
RT "An essential role for p300/CBP in the cellular response to hypoxia",
RL Proc. Natl. Acad. Sci. U.S.A. 93:12963-12973(1996).
RN [5]
RN INTERACTIONS WITH NR3C1.
RX MEDLINE=98250578; PubMed=95069696;
RX Fryer C.J., Archer T.K.;
RT "Chromatin remodelling by the glucocorticoid receptor requires the
RT BRG1 complex",
RL Nature 393:88-91(1998).
RN [6]
RN INTERACTION WITH NCOA6.
RX MEDLINE=20283976; PubMed=10823961;
RX Ko L., Cardona G.R., Chin W.W.;
RT "Thyroid hormone receptor-binding protein, an LXXLL motif-containing
RT protein, functions as a general coactivator",
RL Proc. Natl. Acad. Sci. U.S.A. 97:6212-6217(2000).
RN [7]
RN INTERACTION WITH DTX1.
RX MEDLINE=21576166; PubMed=11564735;
RX Yamamoto N., Yamamoto S.-I., Inagaki F., Kawauchi M., Fukamizu A.,
RA Kishi N., Matsuno K., Nakamura K., Weinmaster G., Okano H.,
RA Nakafuku M.;
RT "Role of Deltex-1 as a transcriptional regulator downstream of the
RT Notch receptor",
RL J. Biol. Chem. 276:45031-45040(2001).
RN [8]
RN PHOSPHORYLATION, AND INTERACTION WITH TCF7L2 AND LEF1.
RX MEDLINE=22450614; PubMed=12446687;
RX Hecht A., Stemmler M.P.;
RT "Identification of a promoter-specific transcriptional activation
RT domain at the C-terminus of the Wnt-effector protein TCF4",
RL J. Biol. Chem. 278:3776-3785(2003).
RN [9]
RN VARIANTS BREAST CANCER PRO-927 AND GLY-1013, VARIANT PANCREATIC CANCER
RP TYR-1650, AND VARIANT COLORECTAL CANCER GLN-2221.
RX MEDLINE=20164332; PubMed=10700188;
RX Gayther S.A., Batley S.J., Linger L., Bannister A., Thorpe K.,
RX Chin S.-F., Daigo Y., Russell P., Wilson A., Sowter H.M.,
RA Delhanty J.D.A., Ponder B.A.J., Kouzarides T., Caldas C.;
RT "Mutations truncating the EP300 acetylase in human cancers",
RL Nat. Genet. 24:300-303(2000).
RN [10]
RN STRUCTURE BY NMR OF 302-418 IN COMPLEX WITH 786-826 OF HIF1A.
RX MEDLINE=21957254; PubMed=11959990;
RX Freedman S.J., Sun Z.Y., Poy F., Kung A.L., Livingston D.M.,
RA Wagner G., Eck M.J.;
RT "Structural basis for recruitment of CBP/p300 by hypoxia-inducible
RT factor-1 alpha.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:5367-5372(2002).
RN [11]
RN FUNCTION: Functions as histone acetyltransferase and regulates
CC transcription via chromatin remodeling. Acetylates all four core
CC histones in nucleosomes. Histone acetylation gives an epigenetic
CC tag for transcriptional activation. Binds to and may be involved
CC in the transforming capacity of the adenovirus E1A protein.
CC [12]
CC CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.
CC [13]
CC SUBUNIT: Interacts with NR3C1, PCAF, NCOA6, LEF1 and TCF7L2. The
CC TAZ-type 1 domain interacts with HIF1A. Probably part of a complex
CC with HIF1A and CREBBP. Interacts with DTX1.
CC [14]
CC SUBCELLULAR LOCATION: Nuclear.
CC [15]
CC PTM: Phosphorylated.
CC [16]
CC DISEASE: Defects in EP300 may play a role in epithelial cancer.
CC [17]
CC SIMILARITY: Contains 1 bromodomain.
CC [18]
CC SIMILARITY: Contains 1 ZAZ-type zinc finger.
CC [19]
CC SIMILARITY: Contains 2 TAZ-type zinc fingers.
CC [20]
CC DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW=http://www.infobiogen.fr/services/Chromancer/Genes/P300ID97.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
RX EMBL; U01877; AAA18639.1; -;
RX PIR; A54277; A54277.
RX PDB; 1L3E; 11-FEB-03.
RX TRANSFAC; T01427; -;
RX Genew; HGNC:3373; EP300.
RX MIM; 602700; -;
RX GO; GO:0008022; P:protein C-terminus binding; TAS.
RX GO; GO:0003713; F:transcription co-activator activity; TAS.
RX GO; GO:0003700; F:transcription factor activity; TAS.
RX GO; GO:0007399; P:neurogenesis; TAS.
RX GO; GO:0007165; P:signal transduction; TAS.
RX InterPro; IPR001487; Bromodomain.
RX InterPro; IPR003101; KIX.
RX InterPro; IPR00197; TAZ finger.
RX InterPro; IPR00433; Znf_ZZ.
RX Pfam; PF00439; Bromodomain; 1.
RX Pfam; PF02172; KIX; 1.
RX Pfam; PF02135; zf-TAZ; 2.
RX Pfam; PF00569; Zf; 1.
RX PRINTS; PR00503; BROMODOMAIN.
RX SMART; SM00297; BROMO; 1.
RX SMART; SM00551; Znf_TAZ; 2.
RX SMART; SM00291; Znf_Zz; 1.
RX PROSITE; PS00633; BROMODOMAIN_1; 1.
RX PROSITE; PS0014; BROMODOMAIN_2; 1.
RX PROSITE; PS0134; ZF_TAZ; 2.
RX PROSITE; PS0135; ZF_ZZ_1; 1.
RX PROSITE; PS0135; ZF_ZZ_2; 1.
RX Transfaser; Transcription regulation; Nuclear protein; Bromodomain;
KW Cell cycle; Zinc-finger; Phosphorylation; Disease mutation;
KW 3D-structure.
FT DOMAIN 11 17 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT ZN_FING 331 417 TAZ-TYPE 1.

FT DOMAIN 1067 1139 BROMODOMAIN.
FT DOMAIN 1572 1818 BINDING REGION FOR E1A ADENOVIRUS.
FT ZN_FING 1684 1707 ZN-TYPE.
FT ZN_FING 1728 1809 TAZ-TYPE 2.
FT DOMAIN 797 800 POLY-SER.
FT DOMAIN 1519 1526 POLY-GLU.
FT DOMAIN 2066 2069 POLY-GLN.
FT DOMAIN 2190 2195 POLY-GLN.
FT VARIANT 827 827 L -> P (in breast cancer).
FT VARIANT 1013 1013 /FTID=VAR_014428.
FT VARIANT 1650 1650 E -> G (in breast cancer).
FT VARIANT 2221 2221 S -> Y (in pancreatic cancer).
FT VARIANT 2221 2221 /FTID=VAR_014430.
FT VARIANT 2221 2221 P -> Q (in colorectal cancer).
FT SEQUENCE 2414 AA; 264143 MW; 68FP909BE4B9D693 CRC64; /FTID=VAR_014431.
Query Match 3.7%; Score 132.5; DB 1; Length 2414;
Best Local Similarity 21.1%; Pred. No. 3.4;
Matches 128; Conservative 75; Mismatches 235; Indels 169; Gaps 31;
QY 4 PREGQSPBPQ-----PWGRLRLGAEEGEPH-----VLL-----RRREWTI 40
Db 300 FNMGOQ-FAPQVQVQGLVTPVAGVSGGTAADPEKRLIQOQLVLLHAHKQEREQAN 358
QY 41 GRRRGCDLSF---PKNLVSGDHRIVDEKSGQVLTEDTSGTVINKLVKKVKTQCP 97
Db 359 GEVRCQNLPHCRTRKXNLNHTHCQ---SGKSCQVA--HCASSRQIISHWKXNTRHDCPV 413
QY 98 -----QTGDIIVLVKNEPEHNVAIVYLSLSEKQGMQESFE--ANKENVFHGTQDSG 150
Db 414 CULPLNAGD-----KKNQOPIITGAPVGLGNFSSLGVGQOQAPNLSTVSDIPPSIERAY 468
QY 151 AGAGRGADPRVPFSPATQCFEE-PQPSTSTDLFTTASASTPSPAGRSRSCSG 209
Db 469 AALGLPYQVQNPQTQVQAKNQOQPGSQGKRPMSNMSA---SPMG-----VNG 518
QY 210 GGGISPKGGPVASDEVSPASALPDR-----KTASPSLEPQD---QEDLEPVKKMR 261
Db 519 GVGQVT-----PSLLSD--SNLHSAINSONPMMSNASVPSLGMPTAAQSPSTTGRKQWH 572
QY 262 GDGDLINGQLLVAQPRNAQTVHEDVRAAGKPD-----KMBETLTCTI--ICODLLH 312
Db 573 EDITQDLRHL-----VHKLVAIFPTPDPAALKDRRMENLVAYARKVSGDMVE 621
QY 313 DCVSLQPMHTFCAACYSGMERSLCPTRCPVERICQKHLNN---LVEAYLIQHDK 369
Db 622 SANNAEYVHLLAEKTYKIQKELEKRRT-----RLQKNLPLNAGMVPVSMNPGPNM 675
QY 370 SRSEEDVQND-----ARNKITQMLQPKVRSFSDSEGSDELLELSDVDSSESDIS 422
Db 676 GQFQPGMTSGPLPDPFDMIRGVSFNQMPRIPTQSGLNQFG-----QMSWA 721
QY 423 QYVVCRCQP--BYRRQAAP-----PHCPAPEGE-----P 451
Db 722 QPPIVPRQTPLQHHGQLAQPGALNPMGYGRMQQPSNQGFPLPOTQFPFSGQMVNTIP 781
QY 452 GAPQALGDAPSTSVSLTAVQDVVCL-----QGSNALCTCFQWPDPRRAREQD 502
Db 782 LAPSS--GAQVSAQAMSSS-----SCPVNSPIMPFGSQGSHIHCPQLQP-----ALHQS 831
QY 503 PRVAPQ 509
Db 832 PVPFPR 838
RESULT 12
RN12 MOUSE STANDARD; PRT; 600 AA.
ID RN12 MOUSE
AC Q9NTV7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

10-OCT-2003 (Rel. 42, Last annotation update)
DE RING finger protein 12 (LIM domain interacting RING finger protein)
GN RNF12 OR RLIM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bach I.;
RT "Opposing developmental functions of positive and negative
coregulators of LIM homeodomain factors.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP FUNCTION.
RA MEDLINE=99364422; PubMed=10431247;
RA Bach I., Rodriguez-Esteban C., Cartiere C., Bhushan A., Krones A.,
Rose D.W., Glass C.K., Andersen B., Izpisua-Belmonte J.-C.,
Rosenfeld M.G.;
RT "RIM inhibits functional activity of LIM homeodomain transcription
factors via recruitment of the histone deacetylase complex.";
RL Nat. Genet. 22:394-399(1999).
CC !- FUNCTION: Acts as a negative co-regulator for LIM homeodomain
transcription factors. Via the recruitment of the SIN3a/histone
deacetylase corepressor complex.
CC !- SUBUNIT: Associates with LIM/homeobox factors.
CC !- SIMILARITY: Contains 1 RING-type zinc finger.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; AF069992; AAD34209.1; -
DR MGD; MGI:1342291; Rnf12.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00519; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Transcription regulation; Zinc-finger.
FT DOMAIN 415 484 SER-RICH.
FT ZN_FING 546 587 RING-TYPE.
FT DOMAIN 447 461 POLY-SER.
FT SEQUENCE 600 AA; 66470 MW; 568299530126E41D CRC64;
Query Match 3.7%; Score 131.5; DB 1; Length 600;
Best Local Similarity 23.4%; Pred. No. 0.7;
Matches 60; Conservative 29; Mismatches 80; Indels 87; Gaps 11;
QY 100 GDVILVYRKNEPEHNVAIVYLSLSEKQGMQESFEANKENVFHGTXTDTSAGAGRGADP 159
Db 407 GELSYFMISDSDFSA-----SVSSRNVERVES-----RNGRSGSGKNSG-- 449
QY 160 RVPSPSPATQVCFEPQFSTSTDLFTTASASTPSPAGRSRSCSGGGGSPKSG 219
Db 450 -----SSSSSSPSSSGSESSESSSKMFESEGGSS-----SG 482
QY 220 PS-----VASDEVSSPASALPDRKTASFSSLEPQDQEDLEFPVKKRGDGLDLNG 270
Db 483 PSRXDGRHRAPTFDE-----SGSLPFFSLAQFFLN--EDDED----- 519
QY 271 QLLVAQPR-RNAQTVHEDVRAAGKPKMVEETLTCTIIC--QDLHDCVSLQPCMHHTCAA 327
Db 520 -----QPRGLTKEQIDNLNRSFGENDALK---TCSVCITEYTEGDKRLKLPCHSEFHVH 571
QY 328 CYSGWMSRSSLCPTCR 343
Db 572 CIDRWLSENSTCPICR 587

RESULT 13	
ID	IGA4_HABIN
STANDARD;	PRT; 1949 AA.
AC	P45386;
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1 protease).
GN	IGA.
OS	Haemophilus influenzae.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC	Pasteurellaceae; Haemophilus.
OX	NCBI_TaxID=727;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=NHTI HK61;
RX	MEDLINE=92234949; PubMed=1373717;
RA	Poulsen K., Reinholdt J., Kilian M.;
RT	"A comparative genetic study of serologically distinct Haemophilus
RT	influenzae type 1 immunoglobulin A1 proteases.";
RL	J. Bacteriol. 174:2913-2921(1992).
CC	- - FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC	PRODUCING INTACT FC AND FAB FRAGMENTS.
CC	- - CATALYTIC ACTIVITY: Cleavage of immunoglobulin A molecules at
CC	certain Pro- -Xaa bonds in the hinge region. No small molecule
CC	substrates are known.
CC	- - SUBCELLULAR LOCATION: Secreted.
CC	- - DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC	SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC	OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC	DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC	SIMILARITY).
CC	- - SIMILARITY: Belongs to peptidase family S6.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M87491; AAA24968.1; -
DR	MEROPS; S06.001; -
DR	InterPro; IPR006315; Autotransport.
DR	InterPro; IPR005546; Autotransporter.
DR	InterPro; IPR000710; Peptidase_S6.
DR	InterPro; IPR004899; Pertactin.
DR	Pfam; PF03797; Autotransporter; 1.
DR	Pfam; PF02395; IGA1; 1.
DR	Pfam; PF03212; Pertactin; 1.
DR	PRINTS; PRO0921; IGASRPRTASE.
DR	TIGRPFAMS; TIGR01414; autotrans_bar1; 1.
KW	Hydrolase, Serine protease; Transmembrane; Zymogen; Signal.
FT	SIGNAL 1 25 POTENTIAL.
FT	CHAIN 26 1021 IMMUNOGLOBULIN A1 PROTEASE.
FT	PROPEP 1022 1849 HELPER PEPTIDE (POTENTIAL).
FT	ACT_SITE 299 299 PROBABLE.
SQ	SEQUENCE 1849 AA; 202957 MW; 79A7D018C7150AEA CRC64;
Query Match 3.7%; Score 131; DB 1; Length 1849;	
Best Local Similarity 19.9%; Pred. No. 3;	
Matches 99; Conservative 67; Mismatches 185; Indels 146; Gaps 22;	
Qy	53 NKLVSQDHCRIVVDEKSG-----QVLEDTSTGTVINKLVKVKQTCPLQTDGVLYVY 107
Db	930 NKSATGNFTLQVAD-KTGEFNNHNLTFDASNA--TENNLEVT-----LANGSVDRGAW 980
Qy	108 RKNPEHNVAI-IYESLSKQCTQE-----SFEANKENVFHGTKDTSGA 151
Db	981 KYKLNVNGRYDLYNPEVKNQVDTTNTTNDIQADAPSAQSNNEIAR----- 1032

RESULT 14	
ID	RN27_MOUSE
STANDARD;	PRT; 551 AA.
AC	Q99PJ2; Q99PQ4;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	RING finger protein 27 (Glioblastoma-expressed ring finger protein)
DE	(Tripartite motif protein 8).
GN	TRIM8 OR RNF27 OR GERP.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Brain;
RX	MEDLINE=20568703; PubMed=1118312;
RA	Vincent S.R., Kwanicka D.A., Fretier P.;
RT	"A novel RING finger-B box-coiled-coil protein, GERP.";
RL	Biochem. Biophys. Res. Commun. 279:482-486(2000).
RN	[2]
RP	SEQUENCE OF 196-463 FROM N.A.
RX	MEDLINE=21231161; PubMed=11331580;
RA	Reymond A., Meroni G., Fantozzi A., Merla G., Cairo S., Luzi L.,
RA	Riganelli D., Zanaria E., Messali S., Cainarca S., Guffanti A.,
RA	Minucci S., Pelicci P.G., Ballabio A.;
RT	"The tripartite motif family identifies cell compartments.";
RL	EMBO J. 20:2140-2151(2001).
CC	- - SUBUNIT: Homodimer (By similarity).
CC	- - SUBCELLULAR LOCATION: Nuclear (By similarity).
CC	- - DEVELOPMENTAL STAGE: At 10.5 and 12.5 dpc, expressed in the
CC	central nervous system. At 14.5 dpc, expressed in the eye (lens
CC	and inner neural layer of the retina), in the primitive glomeruli
CC	of the developing kidney, in the villi of the gut and in the
CC	dorsal root ganglia.
CC	- - DOMAIN: The coiled coil domain is required for homodimerization
CC	(By similarity).
CC	- - SIMILARITY: Contains 1 RING-type zinc finger.
CC	- - SIMILARITY: Contains 2 B box-type zinc fingers.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF281047; AAG53088.1; -.
EMBL; AF220035; AAG53489.1; -.
MGD; MGI:193302; Trm.
GO; GO:0005634; C:nucleus, IDA.
GO; GO:0005515; F:protein binding; IPI.
InterPro; IPR000315; Znf_Box.
InterPro; IPR001841; Znf_ring.
Pfam; PF00643; Zf-B_Box; 1.
Pfam; PF00097; Zf-G3HC4; 1.
SMART; SM00184; RING; 1.
PROSITE; PS01119; ZF_BOX; FALSE_NEG.
PROSITE; PS00518; ZF_RING; 1.
PROSITE; PS50089; ZF_RING; 2; 1.
KW Nuclear protein; Coiled coil; Zinc-finger; Repeat.
FT ZN_FING 15 56 RING-TYPE
FT ZN_FING 92 132 B_BOX-TYPE 1.
FT ZN_FING 140 182 B_BOX-TYPE 2.
FT DOMAIN 181 249 COILED COIL (POTENTIAL).
FT DOMAIN 274 295 COILED COIL (POTENTIAL).
FT CONFLICT 259 259 G -> A (IN REF. 2).
SQ SEQUENCE 551 AA; 61591 MW; CECB2D70C8B130B6 CRC64;

Query Match 3.6%; Score 130.5; DB 1; Length 551;
Best Local Similarity 23.4%; Pred. No. 0.72;
Matches 57; Conservative 21; Mismatches 79; Indels 87; Gaps 11;
QY 299 EETLTCTTCODLLHDCVSLQPMFTCAACY-SGMERSL--QPTCRCPVER---ICKN 352
D5 10 EEELICPICLHVFEVPQL-PCKNFCRCGICGEAWKDSGLVRCPECNQAYNQKPLEKN 68
QY 353 HILNNLVEAYLLQHPDKS-----RSEEDVQSDARNKITQDML 390
D5 69 LKLTNIVEKFNALHVEKPTTALHCVFCRRGPPLEPAQKVLCEAPC---CQSHVQTHLQ 124
QY 391 QPKVRSFSDREGSSEDLLELSDSSDISQPVVWCRCQPEYRQAQPHCPAPEGE 450
D5 125 QPSTARGHL-----LVEADDDVRAWS-----CFQH-NAYRLYHCEAEQ-- 160
QY 451 PGAPALGDAPSTSVLTATVDYVCPQGSALCTCC-----PQMPDPAEAE 498
D5 161 -----VAVCYCCYSGARQGHSCVDVIRNEIRNMLMKQOELEE 202
QY 499 REQD 502
D5 203 REQD 206

RESULT 15
ACIN HUMAN STANDARD; PRT; 1341 AA.
ID ACIN HUMAN 075158; Q9UG91; Q9UKV1; Q9UKV2;
AC Q9UKV3; 075158; Q9UG91; Q9UKV1; Q9UKV2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apoptotic chromatin condensation inducer in the nucleus (Acinus).
GN ACINUS OR KIAA0670.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), PARTIAL SEQUENCE, FUNCTION,
RP AND MUTAGENESIS OF ASP-1093.
RX MEDLINE=99418558; PubMed=10490026;
RA Sahara S., Aoto M., Eguchi Y., Imamoto N., Yoneda Y., Tsujimoto Y.;
RT "Acinus is a caspase-3-activated protein required for apoptotic
chromatin condensation."

Nature 401:168-173(1999).
[2] SEQUENCE FROM N.A. (ISOFORM 4).
RP TISSUE=Uterus;
RC Wambutt R., Heubner D., Mewes H.-W., Gassenhuber J., Wiemann S.;
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[3] SEQUENCE FROM N.A. (ISOFORM 3).
RP TISSUE=Fetal brain;
RC Li W.B., Gruber C., Jessee J., Polayes D.;
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.;
RT "Full-length cDNA libraries and normalization";
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
[4] SEQUENCE OF 56-1341 FROM N.A. (ISOFORM 1).
RP TISSUE=Brain;
RC MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.;
RT "Prediction of the coding sequences of unidentified human genes. X.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro";
RT DNA Res. 5:169-176(1998).
CC -!- FUNCTION: Induces apoptotic chromatin condensation after
activation by CASP3.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=S;
CC IsoId=Q9UKV3-1; Sequences=Displayed;
CC Name=2; Synonyms=S';
CC IsoId=Q9UKV3-2; Sequences=VSP_004025, VSP_004028;
CC Name=3; Synonyms=S;
CC IsoId=Q9UKV3-3; Sequences=VSP_004026, VSP_004029;
CC Name=4;
CC IsoId=Q9UKV3-4; Sequences=VSP_004027;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- PTM: Undergoes proteolytic cleavage; the processed form is active,
contrary to the uncleaved form.
CC -!- SIMILARITY: Contains 1 SAP domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF124726; AAD56724.1; -.
EMBL; AF124727; AAD56725.1; -.
EMBL; AF124728; AAD56726.1; -.
EMBL; AL050382; CAB43681.1; -.
EMBL; BX247975; CAB62309.1; -.
EMBL; AB014570; BAA31645.2; -.
Genew; HGNC:17066; ACINUS.
MIM; 604562; -.
GO; GO:0005634; C:nucleus, IDA.
GO; GO:0016887; F:ATPase activity; NAS.
GO; GO:0019899; F:enzyme binding; NAS.
GO; GO:0003676; F:nucleic acid binding; NAS.
GO; GO:0030263; P:apoptotic chromosome condensation; IDA.
GO; GO:0030218; P:erythrocyte differentiation; IEP.
GO; GO:0045657; P:positive regulation of monocyte differentiation. . . IEP.
InterPro; IPR003034; SAP.
Pfam; PF02037; SAP; 1.
SMART; SM00513; SAP; 1.
PROSITE; PS50800; SAP; 1.
Apoptosis; Nuclear protein; Alternative splicing.
DOMAIN 72 106 SAP.
DOMAIN 142 442 GLU-RICH.
DOMAIN 573 676 SER-RICH.
DOMAIN 1114 1131 PRO-RICH.
DOMAIN 1132 1341 ARG/ASP/GLU/LYS-RICH.

Search completed: May 7, 2004, 14:46:30
Job time : 26.4372 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:40:47 ; Search time 96.0591 Seconds
(without alignments)

2180.991 Million cell updates/sec

Title: US-10-048-046-2

Perfect score: 3585

Sequence: 1 MERPEEGKQSPPPQPGWGRLL.....VKAHAMKFNHICEQTRFKN 664

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3585	100.0	664	Q9NRt4	Q9nr4 homo sapien
2	3504	97.7	652	Q96SL3	Q96sl3 homo sapien
3	3500	97.6	652	Q96EP1	Q96ep1 homo sapien
4	3162.5	88.2	623	Q9NVD5	Q9nvd5 homo sapien
5	2916	81.3	664	Q810L3	Q810l3 mus musculus
6	2891.5	80.7	663	Q8BJZ9	Q8bjz9 mus musculus
7	1678	46.8	306	Q9NT32	Q9nt32 homo sapien
8	896.5	25.0	254	Q8BWH4	Q8bwh4 mus musculus
9	357.5	10.0	473	Q9SX88	Q9sx88 arabidopsis
10	167.5	4.7	540	Q7ZX20	Q7zx20 xenopus lae
11	167	4.7	485	Q803C1	Q803c1 brachydanio
12	160.5	4.5	352	Q9BYE7	Q9bye7 homo sapien
13	160.5	4.5	432	Q8IUD6	Q8iud6 homo sapien
14	160	4.5	1546	Q7Z3L3	Q7z3l3 homo sapien
15	155.5	4.3	733	Q9RI88	Q9rl88 mus musculus
16	154.5	4.3	731	Q8NH2	Q8nh2 mus musculus

17	152	4.2	245	11	Q9CQ80	Q9cqe0 mus musculo
18	151.5	4.2	350	4	Q8N3R7	Q8n3r7 homo sapien
19	150	4.2	513	3	Q9CLX4	Q9clx4 schizosacch
20	149	4.2	209	11	Q99PD2	Q99pd2 rattus norv
21	149	4.2	689	10	Q9FNZ7	Q9fnz7 oryza sativ
22	148.5	4.1	209	11	Q9WTR9	Q9wtr9 mus musculo
23	148	4.1	753	11	Q9D4H7	Q9d4h7 mus musculo
24	146.5	4.1	220	4	Q96SJI	Q96sj1 homo sapien
25	146.5	4.1	277	4	Q96ID9	Q96id9 homo sapien
26	146	4.1	303	4	Q7Z4I9	Q7z4i9 homo sapien
27	145	4.0	504	16	Q87PU0	Q87pu0 vibrio para
28	144.5	4.0	2322	10	Q8SV27	Q8sv27 arabidopsis
29	143.5	4.0	638	11	Q8BX92	Q8bxp2 mus musculo
30	143.5	4.0	700	12	Q7T400	Q7t400 simian herp
31	143.5	4.0	887	11	Q8CFF5	Q8cff5 mus musculo
32	143.5	4.0	1033	11	Q8CGC8	Q8cgc8 mus musculo
33	143.5	4.0	1033	11	Q8OZ37	Q8oz37 mus musculo
34	143	4.0	1102	5	Q9V5P3	Q9v5p3 drosophila
35	143	4.0	1601	5	Q9V6I9	Q9v6i9 drosophila
36	142	4.0	532	3	Q8WZL0	Q8wz10 neurospora
37	142	4.0	781	3	Q9C2M7	Q9c2m7 neurospora
38	142	4.0	818	4	Q8WJ72	Q8wu72 homo sapien
39	141.5	3.9	424	10	Q9C8W4	Q9c8w4 arabidopsis
40	141.5	3.9	444	10	Q9M621	Q9m621 arabidopsis
41	141.5	3.9	448	10	Q8VZ14	Q8vz14 arabidopsis
42	141.5	3.9	628	13	Q7ZVC1	Q7zvcl brachydanio
43	141.5	3.9	652	4	Q86FI5	Q86fi5 homo sapien
44	141.5	3.9	1208	4	Q75162	Q75162 homo sapien
45	140.5	3.9	765	5	Q9UQA3	Q9uaq3 caenorhabdi

ALIGNMENTS

RESULT 1

ID Q9NRt4 PRELIMINARY; PRT; 664 AA.

AC Q9NRt4;

DT 01-OCT-2000 (T-EMBLrel. 15, Created)

DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)

DE 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Cell cycle checkpoint protein CHFR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20388685; PubMed=10935642;

RA Scolnick D.M., Halazonetis T.D.;

RT "Chfr defines a mitotic stress checkpoint that delays entry into metaphase."

RL Nature 406:430-435(2000).

CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

DR EMBL; AF170724; AAF91084.1; -.

DR PDB; 1LGP; 07-AUG-02.

DR PDB; 1LQG; 07-AUG-02.

DR Genew; HGNC:20455; CHFR.

DR GO; GO:0007093; P:mitotic checkpoint; TAS.

DR InterPro; IPR000253; FHA.

DR InterPro; IPR008984; SMAD_FHA.

DR InterPro; IPR001841; Znf_Fing.

DR Pfam; PF00498; FHA; 1.

DR Pfam; PF00097; zf-C3HC4; 1.

DR SMART; SMO0240; FHA; 1.

DR SMART; SMO0184; RING; 1.

DR PROSITE; PSS00006; FHA DOMAIN; 1.

DR PROSITE; PSS00518; ZF_RING_1; 1.

DR PROSITE; PSS0089; ZF_RING_2; 1.

KW Metal-binding; Zinc; Zinc-finger.

SQ SEQUENCE 664 AA; 73386 MW; 141A1E7FEFAB36A2 CRC64;

Query Match

100.0%; Score 3585; DB 4; Length 664;

Best Local Similarity 100.0%; Pred. No. 2e-275;
Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERPEEGKQSPPOPGWRLRLRGAEAGEPHVLLRKREWTIGRRRCDDLSFSPNKLVSGDH 60
DB 1 MERPEEGKQSPPOPGWRLRLRGAEAGEPHVLLRKREWTIGRRRCDDLSFSPNKLVSGDH 60
QY 61 CRIVDEKSGQVLTEDTSTGTVINKLVKKYKQTCPLQTDVILYVYRKNPEHNVAIY 120
DB 61 CRIVDEKSGQVLTEDTSTGTVINKLVKKYKQTCPLQTDVILYVYRKNPEHNVAIY 120
QY 121 ELSSEKQGMTOESFEANKENVEFGTDTSGAGAGRGADPRVPSSPATVCFEEQPST 180
DB 121 ELSSEKQGMTOESFEANKENVEFGTDTSGAGAGRGADPRVPSSPATVCFEEQPST 180
QY 181 TSDLPFTASASSTEPSAGRRSSSCGSGGGISPKGSGPSVASDEVSSFASALPDKRTA 240
DB 181 TSDLPFTASASSTEPSAGRRSSSCGSGGGISPKGSGPSVASDEVSSFASALPDKRTA 240
QY 241 SFSSLEPQDOEDLEPVKKVKGDDLDLNGQLLVAPRRNAQTVDHEDVRAAGKPKDKEE 300
DB 241 SFSSLEPQDOEDLEPVKKVKGDDLDLNGQLLVAPRRNAQTVDHEDVRAAGKPKDKEE 300
QY 301 TLTCIIICODLLHDCVSLQPCMHFTCAACYSYGMWERSLCTCRCPVERICKNHILNLVE 360
DB 301 TLTCIIICODLLHDCVSLQPCMHFTCAACYSYGMWERSLCTCRCPVERICKNHILNLVE 360
QY 361 AYLIQHPDKSRSEEDVQSDARKNITQDMLQPKVRSFSDSESSDLELSDVDSSESD 420
DB 361 AYLIQHPDKSRSEEDVQSDARKNITQDMLQPKVRSFSDSESSDLELSDVDSSESD 420
QY 421 ISQYVVCRCQPEYRQAAQPPHCPAEPGEPGAPQALGAPSTSVLTAVQDYVCPLOQ 480
DB 421 ISQYVVCRCQPEYRQAAQPPHCPAEPGEPGAPQALGAPSTSVLTAVQDYVCPLOQ 480
QY 481 SHALCTCCFPMPDRRAERQDPRVAPQCAVCLQPFCHLYGCTRTGCGCLAPFCELN 540
DB 481 SHALCTCCFPMPDRRAERQDPRVAPQCAVCLQPFCHLYGCTRTGCGCLAPFCELN 540
QY 541 LGDKCLDGVYNNNSYESDILKYLATRLGTWKNMLTESLVALQGVFLLSYRVGTDTVL 600
DB 541 LGDKCLDGVYNNNSYESDILKYLATRLGTWKNMLTESLVALQGVFLLSYRVGTDTVL 600
QY 601 CYCCGLRSFRELTYQYRQNIASELPVATSRPDCYGNCRCTQVKAHMKFNHICEQT 660
DB 601 CYCCGLRSFRELTYQYRQNIASELPVATSRPDCYGNCRCTQVKAHMKFNHICEQT 660
QY 661 RFXN 664
DB 661 RFXN 664

RESULT 2
Q96SL3 PRELIMINARY; PRT; 652 AA.
AC Q96SL3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ14781.
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saiko K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RI "NEDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

EMBL; AK027687; BAB55297.1; -;
DR InterPro; IPR000253; FHA
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001841; Znf_Fing.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 652 AA; 72030 MW; 572F2CE6D1743D80 CRC64;

Query Match 97.7%; Score 3504; DB 4; Length 652;
Best Local Similarity 98.2%; Pred. No. 5.2e-269;
Matches 652; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

QY 1 MERPEEGKQSPPOPGWRLRLRGAEAGEPHVLLRKREWTIGRRRCDDLSFSPNKLVSGDH 60
DB 1 MERPEEGKQSPPOPGWRLRLRGAEAGEPHVLLRKREWTIGRRRCDDLSFSPNKLVSGDH 60
QY 61 CRIVDEKSGQVLTEDTSTGTVINKLVKKYKQTCPLQTDVILYVYRKNPEHNVAIY 120
DB 61 CRIVDEKSGQVLTEDTSTGTVINKLVKKYKQTCPLQTDVILYVYRKNPEHNVAIY 120
QY 121 ELSSEKQGMTOESFEANKENVEFGTDTSGAGAGRGADPRVPSSPATVCFEEQPST 180
DB 121 ELSSEKQGMTOESFEANKENVEFGTDTSGAGAGRGADPRVPSSPATVCFEEQPST 180
QY 181 TSDLPFTASASSTEPSAGRRSSSCGSGGGISPKGSGPSVASDEVSSFASALPDKRTA 240
DB 181 TSDLPFTASASSTEPSAGRRSSSCGSGGGISPKGSGPSVASDEVSSFASALPDKRTA 240
QY 241 SFSSLEPQDOEDLEPVKKVKGDDLDLNGQLLVAPRRNAQTVDHEDVRAAGKPKDKEE 300
DB 241 SFSSLEPQDOEDLEPVKKVKGDDLDLNGQLLVAPRRNAQTVDHEDVRAAGKPKDKEE 300
QY 301 TLTCIIICODLLHDCVSLQPCMHFTCAACYSYGMWERSLCTCRCPVERICKNHILNLVE 360
DB 301 TLTCIIICODLLHDCVSLQPCMHFTCAACYSYGMWERSLCTCRCPVERICKNHILNLVE 360
QY 361 AYLIQHPDKSRSEEDVQSDARKNITQDMLQPKVRSFSDSESSDLELSDVDSSESD 420
DB 361 AYLIQHPDKSRSEEDVQSDARKNITQDMLQPKVRSFSDSESSDLELSDVDSSESD 420
QY 421 ISQYVVCRCQPEYRQAAQPPHCPAEPGEPGAPQALGAPSTSVLTAVQDYVCPLOQ 480
DB 421 ISQYVVCRCQPEYRQAAQPPHCPAEPGEPGAPQALGAPSTSVLTAVQDYVCPLOQ 480
QY 481 SHALCTCCFPMPDRRAERQDPRVAPQCAVCLQPFCHLYGCTRTGCGCLAPFCELN 540
DB 481 SHALCTCCFPMPDRRAERQDPRVAPQCAVCLQPFCHLYGCTRTGCGCLAPFCELN 540
QY 541 LGDKCLDGVYNNNSYESDILKYLATRLGTWKNMLTESLVALQGVFLLSYRVGTDTVL 600
DB 541 LGDKCLDGVYNNNSYESDILKYLATRLGTWKNMLTESLVALQGVFLLSYRVGTDTVL 600
QY 601 CYCCGLRSFRELTYQYRQNIASELPVATSRPDCYGNCRCTQVKAHMKFNHICEQT 660
DB 601 CYCCGLRSFRELTYQYRQNIASELPVATSRPDCYGNCRCTQVKAHMKFNHICEQT 660
QY 661 RFXN 664
DB 649 RFXN 652

RESULT 3
Q96EP1 PRELIMINARY; PRT; 652 AA.
ID Q96EP1
AC Q96EP1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg K.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; BC012072; AAH12072.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD FHA.
DR Pfam; PF00498; ZF_RING_1; 1.
DR SMART; SM00184; Znf_Ring.
DR SMART; SM00240; ZF_RING_2; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00518; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 652 AA; 72058 MW; 1A50773CEA679F07 CRC64;

Query Match 97.6%; Score 3500; DB 4; Length 652;
Best Local Similarity 98.0%; Pred. No. 1.1e-268;
Matches 651; Conservative 0; Mismatches 1; Indels 12; Gaps 1;
QY 1 MERPEEGKQSPPPQWGRLLRLGAEAGEPHVLLRKREWTIGRRGCDLSPFNKLVSGDH 60
DB 1 MERPEEGKQSPPPQWGRLLRLGAEAGEPHVLLRKREWTIGRRGCDLSPFNKLVSGDH 60
QY 61 CRIVVDEKSGQVLTSTSGTVINKLVKKQTCPLOTGDIYLYVRKNEPEHNVAYLY 120
DB 61 CRIVVDEKSGQVLTSTSGTVINKLVKKQTCPLOTGDIYLYVRKNEPEHNVAYLY 120
QY 121 ESLSEKQGMQTESPEANKENVFHGTDTSGAGAGRGADPRVPPSPATQVCFEEQPSTS 180
DB 121 ESLSEKQGMQTESPEANKENVFHGTDTSGAGAGRGADPRVPPSPATQVCFEEQPSTS 180
QY 181 TSDFLFTASSTPEPSAGRRSSSCGGGGISPKGSGPSVASDVSPASALPDKRTA 240
DB 169 TSDFLFTASSTPEPSAGRRSSSCGGGGISPKGSGPSVASDVSPASALPDKRTA 228
QY 241 SFSSLEPDQDLEPVKKKRGDGLDLNGQLLVAQPRNAQTVDHEDVRAAAGKPKMEE 300
DB 229 SFSSLEPDQDLEPVKKKRGDGLDLNGQLLVAQPRNAQTVDHEDVRAAAGKPKMEE 288
QY 301 TLTCIIQDILLHDCVSLQPCWHTFCAACYSGWMSRLCPTCRCPVERICKNHLNLYE 360
DB 289 TLTCIIQDILLHDCVSLQPCWHTFCAACYSGWMSRLCPTCRCPVERICKNHLNLYE 348
QY 361 AYLIQHDPKRSREDDVQSMARKITODMLQPKVRSFSDERGSSEDLLELSDVDSGSD 420
DB 349 AYLIQHDPKRSREDDVQSMARKITODMLQPKVRSFSDERGSSEDLLELSDVDSGSD 408
QY 421 ISQYVYVCRQPEYRQAQPPHCPAEPGPAQALGDPASTSVSLTTAVQDYVCPQOG 480
DB 409 ISQYVYVCRQPEYRQAQPPHCPAEPGPAQALGDPASTSVSLTTAVQDYVCPQOG 468
QY 481 SHALCTCCFQPMDDRRAREODPRVAPQCAVCLQPFCHLYWGCRTGCGCLAFPCFLN 540
DB 469 SHALCTCCFQPMDDRRAREODPRVAPQCAVCLQPFCHLYWGCRTGCGCLAFPCFLN 528
QY 541 LGDKCLDGLVNNNSYESDILKNYLATRGLTWKMLTESLVALQRGVFLSDYRVGTDTVL 600
DB 529 LGDKCLDGLVNNNSYESDILKNYLATRGLTWKMLTESLVALQRGVFLSDYRVGTDTVL 588
QY 601 CYCCGLRSFRELTVQYRONIPASELPVNTSRPDCYWGNCRTQVKAHMKFNHICQOT 660
DB 589 CYCCGLRSFRELTVQYRONIPASELPVNTSRPDCYWGNCRTQVKAHMKFNHICQOT 648

QY 661 RPKN 664
DB 649 RPKN 652

RESULT 4

Q9NVD5
ID Q9NVD5 PRELIMINARY; PRT; 623 AA.
AC Q9NVD5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein FLJ10796.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isoqai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.,
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AK001658; BAA91817.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; ZF_RING_1; 1.
DR SMART; SM00240; ZF_RING_1; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00006; FHA DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 623 AA; 69204 MW; 45186D33DAE52711 CRC64;

Query Match 88.2%; Score 3162.5; DB 4; Length 623;
Best Local Similarity 88.5%; Pred. No. 5.7e-242;
Matches 594; Conservative 5; Mismatches 17; Indels 55; Gaps 2;

QY 1 MERPEEGKQSPPPQWGRLLRLGAEAGEPHVLLRKREWTIGRRGCDLSPFNKLVSGDH 60
DB 1 MERPEEGKQSPPPQWGRLLRLGAEAGEPHVLLRKREWTIGRRGCDLSPFNKLVSGDH 60
QY 61 CRIVVDEKSGQVLTSTSGTVINKLVKKQTCPLOTGDIYLYVRKNEPEHNVAYLY 120
DB 61 CRIVVDEKSGQVLTSTSGTVINKLVKKQTCPLOTGDIYLYVRKNEPEHNVAYLY 120
QY 121 ESLSEKQGMQTESPEANKENVFHGTDTSGAGAGRGADPRVPPSPATQVCFEEQPSTS 180
DB 121 ESLSEKQGMQTESPEANKENVFHGTDTSGAGAGRGADPRVPPSPATQVCFEEQPSTS 180
QY 181 TSDFLFTASSTPEPSAGRRSSSCGGGGISPKGSGPSVASDEVSPASA 233
DB 136 TSDFLFTASSTPEPSAGRRSSSCGGGGISPKGSGPSVASDEVSPASA 192
QY 234 LPDRKTASFSLEPDQDLEPVKKKRGDGLDLNGQLLVAQPRNAQTVDHEDVRAAAG 293
DB 193 LPDRKTASFSLEPDQDLEPVKKKRGDGLDLNGQLLVAQPRNAQTVDHEDVRAAAG 252
QY 294 KPDMEETLTCTICQDILLHDCVSLQPCWHTFCAACYSGWMSRLCPTCRCPVERICKN 353
DB 253 KPDMEETLTCTICQDILLHDCVSLQPCWHTFCAACYSGWMSRLCPTCRCPVERICKN 312
QY 354 ILNNLVEAYLIQHPDKRSREDDVQSMARKITODMLQPKVRSFSDERGSSEDLLELSD 413
DB 313 ILNNLVEAYLIQHPDKRSREDDVQSMARKITODMLQPKVRSFSDERGSSEDLLELSD 372

```
QY 414 VDSBSSDIQPVVYVCRQCEPVRRQAAOPPHCPAPEGEPGAPOALGDAPSTSVSLTTAVQD 473
DB 373 VDSBSSDIQPVVYVCRQCEPVRRQAAOPPHCPAPEGEPGAPOALGDAPSTSVSLTTAVQD 432
QY 474 YVCPLOQSHALCTCCFQPMPPDRRAERQDPRVAPQCAVCLQPPCHLYWGCTRTGCGCL 533
DB 433 YVCPLOQSHALCTCCFQPMPPDRRAERQDPRVAPQCAVCLQPPCHLYWGCTRTGCGCL 492
QY 534 APFELNLGDKCLDGLVNNNSYESDILKNYLATRLGLTKNNLTESLVALQGVFLLSDYR 593
DB 493 APFELNLGDKCLDGLVNNNSYESDILKNYLATRLGLTKNNLTESLVALQGVFLLSDYR 552
QY 594 VTGDTVLCYCCGLSFRLETYQYQNTIPASELPVAVTSRDPCTVWGRNCRTOVKAHAMKF 653
DB 553 VTGDTVLCYCCGLSFRLETYQYQNTIPASELPVAVTSRDPCTVWGRNCRTOVKAHAMKF 612
QY 654 NHICEOTREFKN 664
DB 613 NHICEOTREFKN 623

RESULT 5
Q810L3 PRELIMINARY; PRT; 664 AA.
AC Q810L3 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE RIKEN cDNA 5730484M20 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Limb;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049792; AAA9792.1; -.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 2.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 664 AA; 73871 MW; D651BE3E463DEBB6 CRC64;

Query Match 81.3%; Score 2916; DB 11; Length 664;
Best Local Similarity 82.0%; Pred. No. 2, 1e-22;
Matches 547; Conservative 41; Mismatches 73; Indels 6; Gaps 4;

QY 1 MERPEGKQSPPPQWGRLLRLGAEGBPHVLLRKREWTIGRRGCDLSPSNKLVSGDH 60
DB 1 MELHGEQPPPPQEPQWGRLLRLGAEGBPHVLLRKREWTIGRRGCDLSPSNKLVSGDH 60
QY 61 CEIVVDEKSGVLTEDTSGTVINKLVKKQKCPLOTGDIVLVYKNEPENVAVLY 120
DB 61 CKLTVDLSGVELTEDTSGTVINKLVKKQKCPLOTGDIVLVYKNEPENVAVLY 120
QY 121 ESLSEKQGWTOBSFANKENVFHGTSDTSGAGAGADPRVPPSPATQVCFEPQPS 180
DB 121 ESLSGKQSLTQDSLEANKENMFVTKDCS--GPGQGDPPQVPLLSPMAQTCLBEPQPS 178
QY 181 TSDLPPTASSTEP--SPAGHERSSSCGGGGISPKSGFVSASDEVSSFASALPRK 238
DB 179 TSDLLPTASTSTPELTSAQKSSSGPGNTSISPKGRSLVANGELSLSPVQD-K 237
QY 239 TASFSSLEPDQDELEPVKKQKRGDGLDLNGQLLVQAPRRAQTVHEDVRAAAKPKDKM 298
```

```
DB 238 EASFSLLESKDHEELEPAKXKMGDGLDTNLQLLVSGQRNAQTSSEVDKDAASVKPKDM 297
QY 299 EFTLTCIIICODLLHDCVSLQPCWHTFCAACYSGWMSRSLCPTCRCPVERICKKHILNNL 358
DB 298 EFTLTCIIICODLLHDCVSLQPCWHTFCAACYSGWMSRSLCPTCRCPVERICKKHILNNL 357
QY 359 VEAYLIQHPDKRSSEDEVQSMADARKITQDMLOPKVRRSFDEEGSSDLELSVDSES 418
DB 358 VEAYLIQHPDKRSSEDEVQSMADARKITQDMLOPKVRRSFDEEGSSDLELSVDSES 417
QY 419 SDISOPVYVCRQCEPVRRQAAOPPHCPAPEGEPGAPOAL-GDAPSTSVSLTTAVQDYVCP 477
DB 418 SDISOPVYVCRQCEPVRRQAAVQSLPCVPPESELGATLALGGEAPSTASLPTAAPDYMCP 477
QY 478 LQSHALCTCCFQPMPPDRRAERQDPRVAPQCAVCLQPPCHLYWGCTRTGCGCLAPFC 537
DB 478 LQSHALCTCCFQPMPPDRRAERQDPRVAPQCAVCLQPPCHLYWGCTRTGCGCLAPFC 537
QY 538 ELNLGDKCLDGLVNNNSYESDILKNYLATRLGLTKNNLTESLVALQGVFLLSDYRVTGD 597
DB 538 ELNLGDKCLDGLVNNNSYESDILKNYLATRLGLTKNNLTESLVALQGVFLLSDYRVTGD 597
QY 598 TVLCYCCGLSFRLETYQYQNTIPASELPVAVTSRDPCTVWGRNCRTOVKAHAMKFNHIC 657
DB 598 TVLCYCCGLSFRLETYQYQNTIPASELPVAVTSRDPCTVWGRNCRTOVKAHAMKFNHIC 657
QY 658 EOTREFKN 664
DB 658 EOTREFKN 664

RESULT 6
Q8BJZ9 PRELIMINARY; PRT; 663 AA.
AC Q8BJZ9;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to cell cycle checkpoint protein CHFR.
GN 5730484M20RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK077629; BAC36912.1; -.
DR MGD; MGI:2444898; 5730484M20RIK.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00184; RING; 2.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 663 AA; 73855 MW; 42BD653DCB4ADF45 CRC64;

Query Match 80.7%; Score 2891.5; DB 11; Length 663;
Best Local Similarity 81.7%; Pred. No. 1, 9e-220;
Matches 545; Conservative 41; Mismatches 74; Indels 7; Gaps 5;

QY 1 MERPEGKQSPPPQWGRLLRLGAEGBPHVLLRKREWTIGRRGCDLSPSNKLVSGDH 60
```


Db 1 MELHGEQPPPPQPPWGRLLRLGAEDEPQILLWKREWTIGRRRGCDLSPSPNKLVS GDH 60
QY 61 CRIVVDEKSGQVLTETSTSGTVINKLKVVKQTCPLOTGDVILYVYRKNPEHNVAIY 120
Db 61 CKLVDEISGEVLTETSTNGTVINKLVVVKQTYPVPLHSGDIYLYVYRKNPEHNVAIY 120
QY 121 ELSLSEKQGTQTSFEANKENVFHGTGDTGAGAGRGADPRVPSPATQVCEEPQPGTS 180
Db 121 ELSLSEKQGTQTSFEANKENVFHGTGDTGAGAGRGADPRVPSPATQVCEEPQPGTS 178
QY 181 TSLDFTASASTEP--SPAGRRSSSGSGGGISPKGSGPSVASDEVSSPASALPRK 238
Db 179 TSLDFTASASTEP--SPAGRRSSSGSGGGISPKGSGPSVASDEVSSPASALPRK 237
QY 239 TASFSSLEQDQEDLEPVKKRGDGLDQLNGQLLVAOPRRNAQTTHEDVRAAKCPDKM 298
Db 238 EASFSLSEKQGTQTSFEANKENVFHGTGDTGAGAGRGADPRVPSPATQVCEEPQPGTS 297
QY 299 EETLTCTIIQQDLLHDCVSLQPCMHTECAACYSQWMSRSLCTCPCVVERICKNHLNLL 358
Db 298 EETLTCTIIQQDLLHDCVSLQPCMHTECAACYSQWMSRSLCTCPCVVERICKNHLNLL 357
QY 359 VEAYLIQHPDKSRSEEDVQSMARKITQDMLQPKVRRSFDEBGSSEDLLELSDVDS 418
Db 358 VEAYLIQHPDKSRSEEDVQSMARKITQDMLQPKVRRSFDEBGSSEDLLELSDVDS 417
QY 419 SDISQPVVVCRCQPEYRROAOPHCPAPGEGAPQAL-GDAPSTSVSLTTAVODYVCP 477
Db 418 SDISQPVVVCRCQPEYRROAOPHCPAPGEGAPQAL-GDAPSTSVSLTTAVODYVCP 476
QY 478 LOGSHALCTCCFPMPDRRAEQDPRVAPOQCAVCLQPFCHLYWGCTRTGCGCLAPFC 537
Db 477 LOGSHALCTCCFPMPDRRAEQDPRVAPOQCAVCLQPFCHLYWGCTRTGCGCLAPFC 536
QY 538 EUNLGDCKLDGVLNNNSYSDILKNYLATRGITWKNMLTESLVALQGVFLLSDYRVGTG 597
Db 537 EUNLGDCKLDGVLNNNSYSDILKNYLATRGITWKNMLTESLVALQGVFLLSDYRVGTG 596
QY 598 TVLYCCGGLRSFRELTYQYRQNIIPASELPVAVTSRSDCYWGNCRTOVKAHAMKFNHIC 657
Db 597 TVLYCCGGLRSFRELTYQYRQNIIPASELPVAVTSRSDCYWGNCRTOVKAHAMKFNHIC 656
QY 658 EQTRFKN 664
Db 657 EQTRFKN 663

RESULT 7
Q9NT32
ID Q9NT32 PRELIMINARY; PRT; 306 AA.
AC Q9NT32; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hypothetical protein (Fragment).
GN DKFZP434N2420
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Oosterwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137561; CAB70812.1; -;
DR PIR; T46399; T46399.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 306 AA; 6B50F04601FB2939 CRC64;
Query Match 46.8%; Score 1678; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. NO. 1e-124;

Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 359 VEAYLIQHPDKSRSEEDVQSMARKITQDMLQPKVRRSFDEBGSSEDLLELSDVDS 418
Db 1 VEAYLIQHPDKSRSEEDVQSMARKITQDMLQPKVRRSFDEBGSSEDLLELSDVDS 60
QY 419 SDISQPVVVCRCQPEYRROAOPHCPAPGEGAPQALGDAPSTSVSLTTAVODYVCP 478
Db 61 SDISQPVVVCRCQPEYRROAOPHCPAPGEGAPQALGDAPSTSVSLTTAVODYVCP 120
QY 479 QGSHALCTCCFPMPDRRAEQDPRVAPOQCAVCLQPFCHLYWGCTRTGCGCLAPFC 538
Db 121 QGSHALCTCCFPMPDRRAEQDPRVAPOQCAVCLQPFCHLYWGCTRTGCGCLAPFC 180
QY 539 EUNLGDCKLDGVLNNNSYSDILKNYLATRGITWKNMLTESLVALQGVFLLSDYRVGTG 598
Db 181 EUNLGDCKLDGVLNNNSYSDILKNYLATRGITWKNMLTESLVALQGVFLLSDYRVGTG 240
QY 599 TVLYCCGGLRSFRELTYQYRQNIIPASELPVAVTSRSDCYWGNCRTOVKAHAMKFNHIC 658
Db 241 TVLYCCGGLRSFRELTYQYRQNIIPASELPVAVTSRSDCYWGNCRTOVKAHAMKFNHIC 300
QY 659 EQTRFKN 664
Db 301 EQTRFKN 306
RESULT 8
Q9BWH4
ID Q9BWH4 PRELIMINARY; PRT; 254 AA.
AC Q9BWH4; 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Similar to cell cycle checkpoint protein CHFR (Fragment).
GN 5730484M20RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK052473; BAC35008.1; -;
DR MGD; MGI:2444896; 5730484M20RIK.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00498; FHA; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 1.
FT NON_TER
SQ SEQUENCE 254 AA; 27819 MW; 0340571A2BAEB6EF CRC64;
Query Match 25.0%; Score 896.5; DB 11; Length 254;
Best Local Similarity 70.0%; Pred. No. 6.9e-63;
Matches 180; Conservative 23; Mismatches 49; Indels 5; Gaps 3;
QY 1 MERPEEKQPPPPQPPWGRLLRLGAEDEPQILLWKREWTIGRRRGCDLSPSPNKLVS GDH 60
Db 1 MELHGEQPPPPQPPWGRLLRLGAEDEPQILLWKREWTIGRRRGCDLSPSPNKLVS GDH 60
QY 61 CRIVVDEKSGQVLTETSTSGTVINKLKVVKQTCPLOTGDVILYVYRKNPEHNVAIY 120
Db 61 CKLVDEISGEVLTETSTNGTVINKLVVVKQTYPVPLHSGDIYLYVYRKNPEHNVAIY 120
QY 121 ELSLSEKQGTQTSFEANKENVFHGTGDTGAGAGRGADPRVPSPATQVCEEPQPGTS 180
Db 121 ELSLSEKQGTQTSFEANKENVFHGTGDTGAGAGRGADPRVPSPATQVCEEPQPGTS 180

Db 121 ESLGKQSLTQDSLEANKENMFHVKDCS--GPGQDDDPVLLSPMAQTCLEEPQSTS 178

Qy 181 TSDLEPTASASTEP--SPAGRESSSCGGGGISPKGSPGVASDEVSFASALPDRK 238

Db 179 TSDLLPTASTSTPETSAGKHSSSGPGNTSISPKGRSSIVANGELSSSPFQD-K 237

Qy 239 TASFSSLEPQDCELEP 255

Db 238 EASFSSLEKDHLEP 254

RESULT 9

Q9SX88 PRELIMINARY; PRT; 473 AA.

ID Q9SX88

AC Q9SX88

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE F16N3.15

GN F16N3.15

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ecv. Columbia;

RA Vytotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,

RA Lee J., Liu A., Li J., Kremenetskaia I., Gonzalez A., Altafi H.,

RA Araujo R., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,

RA Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C., Rowley D.,

RA Shinn P., Walker M., Davis R.W., Ecker J.R., Federspiel N.A.,

RA Theologis A.

RT "Arabidopsis thaliana chromosome 1 BAC F16N3 sequence."

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

DR EMBL; AC007519; AAD46042.1; -

DR PIR; C96516; C96516.

DR InterPro; IPR001841; Znf_ring.

DR Pfam; PF00097; zf-C3HC4; 1.

DR SMART; SMO0184; RING; 1.

DR PROSITE; PS00089; ZF_RING_2; 1.

KW Metal-binding; Zinc; Zinc-finger.

SQ SEQUENCE 473 AA; 53874 MW; D17217493A81C991 CRC64;

Query Match 10.0%; Score 357.5; DB 10; Length 473;

Best Local Similarity 23.6%; Pred. No. 8.4e-20;

Matches 130; Conservative 65; Mismatches 201; Indels 155; Gaps 20;

Qy 177 PSTSTSLPTASASTEPS-----PAGRESS 204

Db 7 PMTTAE---TGSSGSKPSDDAWAKLVPLDTRFSDIIRCNMVICSEIKPSLEKHE 63

Qy 205 SCGSGGGISPKGSPGVASDEVSSFASALPDRKTASFSLFPQDQEDLEPVKKMRGDP 264

Db 64 WCR-----ITKNLGQSSATHNKSDAILVDK-----AVVPK-----DG 97

Qy 265 DLD-LNGQLLVAQPRNAQTVHEDVRAAGKPKMEETL-----TCITCQDILLHDC 314

Db 98 AVDLISGEIVPGEEGQYQYFTIMPA--PESRTQLIQSIDPBAKCSICLNIWHDV 155

Qy 315 VSLQPCMHTECAACYSQWMERSS-----LCPTCRCPVERICKNHLNINLVEAYLIQHPD 368

Db 156 VTAAPCLHNFCHGCFSEWRSSEKHKHVLCPQRTTVQVGVGNFLKNIQEIILKVDAA 215

Qy 369 KSRSEEDVQSDARKKITQDML--OPKVRFSFDEGSSDLELSDVDSESSDISQPY 425

Db 216 LRPAEDIAVDSASIQSLNLTGSKRKLNPAPTHEERSLR----- 261

Qy 426 VVCRQCFEYRQAAQPHCPGEPGAPQALGAPSTSVLTAVQDYVCPQGHALC 485

Db 262 ---QCPO-----CVANIG-----GYRCEHGAHLQC 284

Qy 486 TCCQPMPPDRARRERQDRVAPQCCAVCVOPFCHLLWGTCTGCGYCLAPPCELN-----L 541

Db 285 HLCQGMFFFRANLQ-----VPLHCKGCDRFFCGAYWS--SENVTQGVSGVVCVRETFPI 337

Qy 542 GDKCLDGV-----LNNNSYESDILKNYLATRGLTWKMLTESL-----VALQGVFLLSDY 592

Db 338 SERTITRIPETHEMAREHODITQRCIAHWEKTVPDVAEMLELFNNRIDSRRPLNHA 397

Qy 593 R-VTGDIVLCYCCGLRSFRELTYQYRQNTIPASLPVAVTSRDPDQWGRNCRTO-VKAAHA 650

Db 398 ETITASTHVCNDCYDKLVGLVFLYFRITLPRNHLPAVAAREDCWYGYACRTQHNEHDA 457

Qy 651 MKNHICEQTR 661

Db 458 RKRNEVCRPTR 468

RESULT 10

Q7ZX20 PRELIMINARY; PRT; 540 AA.

ID Q7ZX20

AC Q7ZX20

DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Similar to ring finger protein (C3HC4 type) 8.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodidae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA Klein S., Strausberg R.,

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC046256; AAH46256.1; -

DR InterPro; IPR000253; FHA.

DR InterPro; IPR008984; SMAD FHA.

DR InterPro; IPR001841; Znf_ring.

DR Pfam; PF00498; FHA; 1.

DR Pfam; PF00097; zf-C3HC4; 1.

DR SMART; SMO0240; FHA; 1.

DR SMART; SMO0184; RING; 1.

DR PROSITE; PS00006; FHA DOMAIN; 1.

DR PROSITE; PS00518; ZF_RING_1; 1.

DR PROSITE; PS50089; ZF_RING_2; 1.

SQ SEQUENCE 540 AA; 51925 MW; A504B796022A1D94 CRC64;

Query Match 4.7%; Score 167.5; DB 13; Length 540;

Best Local Similarity 18.8%; Pred. No. 0.00012;

Matches 101; Conservative 84; Mismatches 216; Indels 135; Gaps 20;

Qy 5 EEGKQPPPPQPGWRLLRLGAEERGEPEHLLRKEWTIGRRRG-----CDLSFFSNKL 55

Db 3 DEG-----PGCMW--CLRRCGRTEDLLLPDGEVTLGRGLGVTYQLKPTLCLP-----M 50

Qy 56 VSDGHCRIVVDKSGQVTLDET-STSGTVINKLVKKQTCPLQTDVLYL-VYRKEPE 113

Db 51 ISTHC-LFKQNTGGVTVTDNKSNGVWRNKREREPHKAAYTLSEGLIQLGPPPNMES 109

Qy 114 HNVAIY--ESLSEKQGM-----TOESFEANKENVFHGTQDTSGAGAGRADPRVPPSSPA 167

Db 110 AEPEYMLVREHLEKLSGLIRPLPKTKATRTKRKFTSEDTDASGNEGSPNFSIPKPYRV 169

Qy 168 TOYCFSEPPQSTSTSLP--PTASASSTEPSPAGERSSSCGSGGGISP----- 215

Db 170 SRE-DEDSAKSHTTULYKQPT-----VEPTASGTESLNDSVEAEVAPTOOQCRSTLQ 223

Qy 216 -----KSGSP----- 220

Db 224 LSRVQTMESIRRLNVQMOKQMEMQEKLSPLENOVGANSVLAQVKELRALHNLHLSNEQ 283

```

273 -----EMODLQTLALRGQLRSQSQBALRMETLEKSFCESEERLETEKAQONEVGLUK 327
QY 271 QL-----LVAQPRRNAQTVHEDVRAAGK-----PDKME 299
Db 328 QLEALKHEKRKVIIBELKHANQGKFEVLQADKSELVTKKEEKAKAKQAEVVTQWTEVL 387
QY 300 ETLTICIQDILLHDCVSLQPMHTFCAACYSGHMWESSLCPTCRCPFEVERICKXHNILNV 359
Db 388 SLSQCSIGSELFIEAVTLN-CAHSFQCHCISEWRNKKDCPMQWNIITSQTSRLVLDNCI 446
QY 360 RAYLIQHPDKRSBEDVQSDMANKITQDMLQPKVARSFSDREGS 405
Db 447 D-----RWYENI-SADMEERLVLINERKGRSKSESIS 481

```

RESULT 12

Q9BYE7	PRELIMINARY;	PRT;	352 AA.
ID	Q9BYE7		
AC	Q9BYE7;		
DT	01-JUN-2001 (TrEMBLrel. 17, Created)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	HMBLR protein.		
DE	HMBLR protein.		
GN	HMBLR.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP	TISSUE=Testis;		
RC	Takahashi N., Akasaka T., Nakayama T., Koseki H., Koga H.;		
RA	"identification and characterization of hMBLR."		
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.		

```
DR STARI; SW00184; RING; 1.
DR PROSITE; PS00282; KAZAL; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
```

Query Match	4.5%;	Score	160.5;	DB	4;	Length	352;
Best Local Similarity	21.4%;	Pred.	No. 0.00023;				
Matches	62;	Conservative	35;	Missmatches	128;	Indels	65;
						Gaps	7;
QY	139	ENVFHGKDTISGAGRGADPRVPISSPATQVCFEEOPOSTSTSDLPFTASASSTESPA	198				
DB	2	EGVAVTAGVGAAKTEGAAALPPP					
QY	199	GRSSSCGGGGGIGSPKGGSPVASDEVSSFASALPDKTKASTSSLEPQDQDLEPVKK	258				
DB	51	SETGAPG					
QY	259	KMRGDGDLNLGOLLYVAQPRRNAQTVDHVRAAAKPKMKEET	306				
DB	91						
QY	307	QODLHPDVSLOPMHTFCAACYSYGMWSSSLCPTCRC	360				
DB	139	CKGLIDATITTECLTFFCKSIIVRHFYNSRCPKNTVHGTOTPLYNIRLDROLQDIV	198				

QY 361 AVLIQHPDKSRSEEDVQSDARNKITQDMLQPKVRSFSDSEGSSEDLLE 410
Db 199 KLVINLEER-----EKKQMHDFYKRGLEVPKPAVPQVPSSKGRSKVLE 244

RESULT 13
Q81UD6
ID Q81UD6 PRELIMINARY; PRT; 432 AA.
AC Q81UD6
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN MGC13061.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21375674; PubMed=11468690;
RA Jenne D.E., Tinschert S., Reimann H., Lasinger W., Thiel G.,
RA Hameister H., Kehrer-Sawatzki H.;
RT "Molecular characterization and gene content of breakpoint boundaries
in NF1 patients with long range deletions.";
RL Am. J. Hum. Genet. 69:516-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Jenne D.E., Tinschert S., Stegmann E., Reimann H., Nuernberg P.,
RA Horn D., Naumann I., Buske A., Thiel G.;
RT "A common set of at least 11 functional genes is lost in the majority
of NF1 patients with gross deletions.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Jenne D.E., Dorschner M.O., Stephens K., Hameister H., Tinschert S.,
RA Kehrer-Sawatzki H.;
RT "Complete physical map and gene repertoire around the NF1 gene in man
and mice.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ496729; CAD43140.1; -
DR Genew; HGNC:21158; RNF135.
DR InterPro; IPR001870; B302.
DR InterPro; IPR003877; SPRY receptor.
DR InterPro; IPR001841; Znf_Fing.
DR Pfam; PF00622; SPRY; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00449; SPRY; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 432 AA; 47873 MW; A6756ECP81A6A2A6 CRC64;

Query Match 4.5%; Score 160.5; DB 4; Length 432;
Best Local Similarity 21.0%; Pred. No. 0.00031;
Matches 85; Conservative 45; Mismatches 158; Indels 117; Gaps 16;

QY 299 EBITLTCIICQLHDCVSLQPCGHTTCAACYSG-WMERSS---LCPTCR---CPVERICK 351
Db 16 EDDLGCICQLL-DWPAFLPCGHSFCHCLALWGARDARRWACFTCRQGAQPHLRK 74

QY 352 NHIINLIVEAY-----LIQHPDKSRSEEDVQSDARNKITQ- 387
Db 75 NTLQLDLADKYRAAREIQAGSDPAHCPCPGSSSLSSAARPRRRPELQKVAVERKITEV 134

QY 388 -----DMLQPKVRSFSDSEGSSEDLLESDVDSSESDIS----- 422
Db 135 AQELTELVELHVDIVRSLOQRFPLESGPDNELSLGKAFSSGVDLSMASPKLVTSDTAA 194

QY 423 -----QPVVVCRQCPEYRRQ-----AAQPPHCPAPEGEPGAPQALGDA 460
Db 195 GKTFDILHLEETQKLESVTWKEAPEAQMQGELLEAPSSSSCPLPDQSHPALREASRF 254

QY 461 PSTSVSTTTAVQDVVCPLOQS--HALCTCCFPMPDR-RAERQDDPRVAPQCAVCLQPP 517
Db 255 AQWAIHTFNLKSLCSLEGSKDSRTVTYSHRQPFYWCERFSTQVL---CSQALSSG 311

QY 518 CHLYWGCTRTGCGYGLAPFCFLNLGDKLDGVNLNNNSYESDIILKNYLATRLGLTWK----- 572
Db 312 KH-YWEVDTRNC-----SHWAVGVASWMSRQDVLGRTMDSCCVWKGTSQ 357

QY 573 ---NMLTESLVALQR-----GVFL-----LSDYRVVTGDTVLCYCC 604
Db 358 SAMHWKETVLGSDRDPGVGILNLBEGKLAIFYVDNQEKLLYEC 402

RESULT 14
Q723L3
ID Q723L3 PRELIMINARY; PRT; 1546 AA.
AC Q723L3
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein DKFZP686H0793 (Fragment).
GN DKFZP686H0793.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human retina;
RA Wamburt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Pobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RW EMBL; BX537764; CAD97829.1; -
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 1546 AA; 170918 MW; 0702441126BAC818 CRC64;

Query Match 4.5%; Score 160; DB 4; Length 1546;
Best Local Similarity 21.1%; Pred. No. 0.0019;
Matches 109; Conservative 74; Mismatches 182; Indels 152; Gaps 24;

QY 51 PSNKLVSQDHCRIVVDKSGQVLTEDTSTGTVINKLVKVKKQTCPLQGTGDIYLVTRKN 110
Db 68 PSHGRVAGFEVQSLQD-AGNQLGMEDTSLSSGML-----TQNTNVPILG----- 111

QY 111 EPEHNVAIYVESLSEKQMGTOESFEA-NKENVFGTKDTSGAGAGRGADPRVPSSPATQ 169
Db 112 -----VDVAISQGITLPSLESFHLNTHIGKGLHATGSKRKKKTLRP-GFVTQ 160

QY 170 -----VCFEPQPSSTSDLPFTASASTEPSPAGHERSSSCSGGGGSPKSGSPSVA 223
Db 161 EDRCDHLTKEP-----FSGEPSEVKEE-----GGKPMNSEGEIPLSLP 200

QY 224 SDVSSFSASALPDRKT-----ASFSLSLPQDOEDLEPV 256
Db 201 SGSQSAKPVSQPRKSTQPDVCASPOEKPLRTLPHQPEEIEDGGLFIPMBEQDNESEKR 260

QY 257 KKKVRG-----DG-----DLDLNGQLIVAQPRRNAQTVDHEDVRAAAGKPKDMBETL 302
Db 261 RKKKGTGKRKEDGRQSGTTLAYDLKLDML-----DRTLEDGAK-----QHNL 303

QY 303 TCITCQDLHLDCVSLQPCGHTFCAA-----CYSQWVERSSI-----CPTCR-C 344
Db 304 TAVNVRNLTVEITHNEHVAMKAAISETEDMPFEPMTRKSKLKEVVEKGVVPTWNIS 363

QY 345 PVERICKHIL-NMLVAYLIQHDPKSRSEEDVQSDARNKIT--QDMLQPKVRSFSD 401
Db 364 PIKK--ANEIKPQFVDIHL---BEDDSSDEYQPDDEEDETAEESILLESDESTASSP 418

QY 402 EGSSEDLI-----ELSDVDSSESDISQPVVCRQCPEYRRQAAQPHCPAPEGEPGAPQAL 457
Db 419 RGAKSKRLRQSEMETDESGILSEAKV--TAPAIRHISAE-----VVPMPGPPPPPKPK 472

Db 207 FKLDHSVSSTNGHRWQIFQDILL-----GTDQDNLDLANVNLMLLELLVQKKQL 254
 QY 416 -----SESSDISQPYVVCRCQ 431
 Db 255 EAESHAQQLIMEFLKVARNRKRELEQIOKELSVLEEDIKRVEEMSGLYSPVSEDSVT 314
 QY 432 PEYRRQAQPPHC-----PAPEGEFGAPQAL-----GDAPSTSVSLTTAVQDYVCP 477
 Db 315 PQF--EAPSPSHSSIIDSTEYSQPPGSGTSQTKQPWYNSTLASRRKRLTAHPED----- 368
 QY 478 LQSHALCTCCFQPMPPDRRAERBQDPRVAPQ--QCAVCLQPFCHLYWCTTGCGCLAP 535
 Db 369 -----LEQCYFS--TRMSRISDDSRASQDDEFQCLSKF-----TR---YNSVRP 409
 QY 536 FCELNIGDKCLDGVNLNNNSYESDILKNYLATRGLTWKQMLTESLVALQRGVFLLSDYRVT 595
 Db 410 LATLSYASDLNGSSIVSSIEFDRCDDYFAIAGVTKIKVYEYGTVIDAV-----DIHYP 465
 QY 596 GDTVLIC-----YCCGLRSFRELTYQYRONIPAS 623
 Db 466 ENEMTCNSKISCISWSS-----YHKULLAS 490

Search completed: May 7, 2004, 14:49:01
 Job time : 100.059 secs

QY 458 GDAPSTSVSLTTAVQDYVVCPLQSHALCTCCFQPMFD 494
 Db 473 QTRDSTFMEXLHVADEEL-----ASSPVCMDSFQPMDD 505
 RESULT 15
 Q9RIAS
 ID Q9RIAS PRELIMINARY; PRT; 733 AA.
 AC Q9RIAS, 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Constitutive photomorphogenic protein.
 GN COP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 EX MEDLINE=93324249; PubMed=10395541;
 RA Wang H., Kang D., Deng X.W., Wei N.;
 RT "Evidence for functional conservation of a mammalian homologue of the
 RL light-responsive plant protein COP1.";
 RL Curr. Biol. 9:711-714(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Yi C., Deng X.-W.;
 RL "The study of mammalian COP1s.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMEL; AF151110; AAD51094.2; -;
 DR MGD; MGI:1347046; Cop1.
 DR InterPro; IPR001680; WD40.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00400; WD40; 6.
 DR Pfam; PF00097; zf-C3HG4; 1.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00184; RING; 1.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS00678; WD_REPEATS_1; 1.
 DR PROSITE; PS00082; WD_REPEATS_2; 2.
 DR PROSITE; PS00294; WD_REPEATS_REGION; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Metal-binding; Repeat; WD repeat; Zinc; Zinc-finger.
 SQ SEQUENCE 733 AA; 80441 MW; 894AEA12BACC737 CRC64;
 Query Match 4.3%; Score 155.5; DB 11; Length 733;
 Best Local Similarity 21.3%; Pred. No. 0.0016;
 Matches 122; Conservative 53; Mismatches 208; Indels 189; Gaps 24;
 QY 153 AGRGADPRVPSPATQVCFEEBPQSTSTDLPTASASTEPSPAGRERSSCGSGGG 212
 Db 7 AGSGAGTSPGSSAASV-----TSAS-----SSLSSSPSPSVAASATLVSGGV 52
 QY 213 ISPKGSG-----PSVASDEVSFASALPDRKTASFSSLEPQDQEDLEFVKKMRGDGD 265
 Db 53 APAAGSGGLGPGRPVLVAASVSSASA-----GGAVSAGQSLSCAARPSAGVG 103
 QY 266 LDINGQLLVAPRNAGTV-----HEDVRAAGKPKMEETLTCTICQDLLHDCVS 316
 Db 104 ---SSSLGSSSRKRLPLVPLCNGLNLSYED-----KSNDFVCPICPDIMEEAY- 149
 QY 317 LQPCMHFTCAACYSYGMWERSLQTCRCPCVERICKHILNNLVEAYLIQHPDKSRSEEDV 376
 Db 150 MTKGHSFCYKCTHQSLSDNNRCFKKNYVDNI--DHLFENFLVNELIL-KQKQFEKR 206
 QY 377 QSMQD-----ARNKIQMQLQPKVRRSPSDESGSLLELSDVD----- 415

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:36:57 ; Search time 14.9546 Seconds
(without alignments)
1379.240 Million cell updates/sec

Title: US-10-048-046-2_COPY_31_103
Perfect score: 378
Sequence: 1 VLLRKEWTIGRRGCDLSF.....INKLKVKKQTCPLQTGQDVI 73

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	378	100.0	426	AAU15856	AAU15856 Human nov
2	378	100.0	426	ABU54925	ABU54925 Human nov
3	378	100.0	623	ABU93182	ABU93182 Human pro
4	378	100.0	623	ABB97233	ABB97233 Novel hum
5	378	100.0	652	ABU93168	ABU93168 Human pro
6	378	100.0	664	ABU83843	ABU83843 Amino aci
7	378	100.0	664	ABU20219	ABU20219 Human Chf
8	304	80.4	92	AAU16156	AAU16156 Human nov
9	304	80.4	92	ABU55225	ABU55225 Human nov
10	275	72.8	99	AAU16157	AAU16157 Human nov
11	275	72.8	99	ABU16576	ABU16576 Human nov
12	275	72.8	99	ABU55645	ABU55645 Human nov
13	275	72.8	99	ABU55226	ABU55226 Human nov
14	107.5	28.4	776	ABU73894	ABU73894 Yeast MEC
15	107.5	28.4	821	AAW28664	AAW28664 Yeast che
16	107.5	28.4	821	ABP54942	ABP54942 Saccharom
17	107.5	28.4	821	ABR53872	ABR53872 Protein s
18	107.5	28.4	821	ABU61612	ABU61612 Yeast SPK
19	99	26.2	699	ABU03445	ABU03445 Candida a
20	99	26.2	699	ABP73798	ABP73798 Candida a
21	86	22.8	176	ABP65604	ABP65604 Bifidobac
22	79	20.9	345	ABR59714	ABR59714 Human tra
23	78.5	20.8	175	ABP64143	ABP64143 Human ORF
24	78.5	20.8	244	AAU16158	AAU16158 Human nov
25	78.5	20.8	244	ABU55227	ABU55227 Human nov

ALIGNMENTS

RESULT 1					
AAU15856					
ID	AAU15856	standard; protein; 426 AA.			
XX	AAU15856;				
AC	AAU15856;				
XX					
DT	07-NOV-2001	(first entry)			
XX					
DE	Human novel secreted protein, Seq ID 809.				
XX					
KW	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;				
KW	cardiac; vasotropic; cerebroprotective; neurotropic; neuroprotective;				
KW	antibacterial; virucide; fungicide; ophthalmological; vulnerary;				
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;				
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;				
KW	cerebral ischaemia; angiogenesis; nervous system disorder;				
KW	Alzheimer's disease; infection; ocular disorder; corneal infection;				
KW	wound healing; epithelial cell proliferation; skin ageing; food additive;				
KW	preservative; antiproliferative.				
XX					
OS	Homo sapiens.				
XX					
PN	WO200155322-A2.				
XX					
PD	02-AUG-2001.				
XX					
PF	17-JAN-2001; 2001WO-US001341.				
XX					
PR	31-JAN-2000; 2000US-0179065P.				
PR	04-FEB-2000; 2000US-0180628P.				
PR	24-FEB-2000; 2000US-0184664P.				
PR	02-MAR-2000; 2000US-0186350P.				
PR	16-MAR-2000; 2000US-0189874P.				
PR	17-MAR-2000; 2000US-0190076P.				
PR	18-APR-2000; 2000US-0198123P.				
PR	19-MAY-2000; 2000US-0205515P.				
PR	07-JUN-2000; 2000US-0209467P.				
PR	28-JUN-2000; 2000US-0214886P.				
PR	30-JUN-2000; 2000US-0215135P.				
PR	07-JUL-2000; 2000US-0216647P.				
PR	07-JUL-2000; 2000US-0216880P.				
PR	11-JUL-2000; 2000US-0217487P.				
PR	11-JUL-2000; 2000US-0217496P.				
PR	14-JUL-2000; 2000US-0218290P.				
PR	26-JUL-2000; 2000US-0220963P.				
PR	26-JUL-2000; 2000US-0220964P.				
PR	14-AUG-2000; 2000US-0224518P.				
PR	14-AUG-2000; 2000US-0224519P.				

Ab47789 Protein x
AB977255 Selected
ABJ11126 Yeast sel
ABJ11042 Yeast sel
ABU36268 Protein e
ABU17261 Protein e
ADA36875 Acinetoba
AAB92565 Human pro
AAY51669 Human nib
AAY32373 Nibrin. e
AG89798 C glutami
AG01640 Human sec
AAW95501 B. subtil
AAW69742 SAPAP2 pr
AAW69740 SAPAP1 pr
ABB63213 Drosophil
ABB52816 Escherich
ABU05697 M. tuberc
AG81105 Mycobacte
ABB11806 Human sec

26 76.5 20.2 517 5 AAB47789
27 72 19.0 199 5 ABG77255
28 72 19.0 199 5 ABJ11126
29 72 19.0 350 5 ABJ11042
30 71 18.8 409 6 ABU36268
31 70.5 18.7 874 6 ABU17261
32 70.5 18.7 877 6 ADA36875
33 68 18.0 399 4 AAB92565
34 68 18.0 754 3 AAY51669
35 68 18.0 754 3 AAY32373
36 66.5 17.6 288 4 AG89798
37 66 17.5 131 3 AG01640
38 66 17.5 288 2 AAW95501
39 66 17.5 992 2 AAW69742
40 66 17.5 992 2 AAW69740
41 65 17.2 945 4 ABB63213
42 63 16.7 138 4 ABB52816
43 62.5 16.5 527 5 ABU05697
44 62.5 16.5 528 4 AG81105
45 62 16.4 1724 4 ABB11806

PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-022182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0236327P.
PR 28-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241121P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-488783/53.
XX N-PSDB; AAS25843.
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
PS Claim 11; SEQ ID NO 809; 980pp; English.
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cardiac
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angioneu-
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,

CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Query Match 100.0%; Score 378; DB 4; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.5e-40;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLRKEWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 60
Db 33 VLLRKEWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 92
QY 61 KKQTCPLQTGDVI 73
Db 93 KKQTCPLQTGDVI 105
RESULT 2
ABU54925
ID ABU54925 standard; protein; 426 AA.
XX AC ABU54925;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polypeptide #12.
XX KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX OS Homo sapiens.
XX PN US2002132753-A1.
XX PD 19-SEP-2002.
XX PF 17-JAN-2001; 2001US-00764864.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218390P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226858P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236388P.
PR 29-SEP-2000; 2000US-0236399P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX PI Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-147444/14.
DR N-PSDB; ABX73184.
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
XX inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX Claim 11; SEQ ID NO 809; 402pp; English.
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX Sequence 426 AA;
SQ Query Match 100.0%; Score 378; DB 6; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.5e-40;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLRKEWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 60
Db 33 VLLRKEWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSGTVINKLVV 92
QY 61 KKQTCPLQTGDVI 73
Db 93 KKQTCPLQTGDVI 105

```
RESULT 3
AAB93182
ID AAB93182 standard; protein; 623 AA.
XX
AC AAB93182;
XX
XX 26-JUN-2001 (first entry)
DT
DE Human protein sequence SEQ ID NO:12128.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
OS
XX EP1074617-A2.
PN
XX 07-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
PR
XX 29-AUG-1999; 99JP-00300253.
PR
XX 11-JAN-2000; 2000JP-00118776.
PR
XX 02-MAY-2000; 2000JP-00183767.
PR
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
DR
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 12128; 2537bp + Sequence Listing; English.
PS
XX
XX The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dr primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides, and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH0166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
XX Sequence 623 AA;
XX
Query Match 100.0%; Score 378; DB 4; Length 623;
Best Local Similarity 100.0%; Pred. No. 2.5e-40;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIWVDEKSGQVLTSTSGTIVNKLKV 60
DB 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIWVDEKSGQVLTSTSGTIVNKLKV 90
XX
RESULT 5
AAB93168
ID AAB93168 standard; protein; 652 AA.
XX
QY 61 KKQTCPLQTGDVI 73
DB 91 KKQTCPLQTGDVI 103
XX
RESULT 4
AAB97233
ID AAB97233 standard; protein; 623 AA.
XX
AC AAB97233;
XX
XX 27-JUN-2002 (first entry)
DT
DE Novel human protein SEQ ID NO: 501.
XX
XX Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
XX antiinfertility; cerebroprotective; cytosatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag.
XX
XX Homo sapiens.
OS
XX WO200222660-A2.
PN
XX 21-MAR-2002.
PD
XX
XX 10-SEP-2001; 2001WO-US026015.
XX
XX 11-SEP-2000; 2000US-00659671.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-292408/33.
DR
XX N-PSDB; AABN32419.
XX
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
XX
XX Example 2; SEQ ID NO 501; 509pp; English.
XX
XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate haemostasis or thrombolysis e.g. to treat
CC infertility, to regulate cancer, to treat inflammatory conditions e.g.
CC and cancer, to screen for drugs, to treat nervous system disorders e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention
XX
XX Sequence 623 AA;
XX
Query Match 100.0%; Score 378; DB 5; Length 623;
Best Local Similarity 100.0%; Pred. No. 2.5e-40;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIWVDEKSGQVLTSTSGTIVNKLKV 60
DB 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIWVDEKSGQVLTSTSGTIVNKLKV 90
XX
QY 61 KKQTCPLQTGDVI 73
DB 91 KKQTCPLQTGDVI 103
XX
RESULT 5
AAB93168
ID AAB93168 standard; protein; 652 AA.
XX
```

AC AAB93168;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12100.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
XX
PR 27-AUG-1999; 99JP-00300253.
XX
PR 11-JAN-2000; 2000JP-00118776.
XX
PR 02-MAY-2000; 2000JP-00183767.
XX
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI: 2001-318749/34.
XX
XX Primer sets for synthesising polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
XX Claim 8; SEQ ID NO 12100; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC polynucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX Sequence 652 AA;
SQ
Query Match 100.0%; Score 378; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 2.6e-40;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTEDTSGTVINKLVV 60
DB 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTEDTSGTVINKLVV 90
QY 61 KKQTCPLQTGDVI 73
DB 91 KKQTCPLQTGDVI 103

RESULT 6
AAB83843
ID AAB83843 standard; protein; 664 AA.
XX
AC AAB83843;
XX
XX 22-AUG-2001 (first entry)
DT
XX Amino acid sequence of a human ring finger protein designated FHAR1.
DE
XX FHAR1; RING finger protein; cancer; vaccine.
KW
XX Homo sapiens.
OS
XX WO200142430-A1.
PN
XX 14-JUN-2001.
PD
XX 07-DEC-2000; 2000WO-US033094.
PF
XX 08-DEC-1999; 99US-00456876.
PR
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA
XX Zhou B, Zhu Y, Chaturvedi P, Hurle MR, Li X;
PI
XX WPI: 2001-381563/40.
DR N-PSDB; AAF89709.
XX
XX New FHAR1 polypeptide, a member of the RING finger protein family for
PT diagnosing and treating cancer, and for use in anti-cancer vaccines.
PT
XX Claim 1; Page 19; 28pp; English.
PS
XX The present sequence represents a FHAR1 polypeptide, which is a member of
CC the RING finger protein family. FHAR1 is useful in the treatment of
CC cancer, and as a vaccine for inducing an immunological response in a
CC mammal. FHAR1 polynucleotides may also be used as a diagnostic reagent
CC through detection of mutations in the associated gene, and for chromosome
CC localization studies, and tissue expression studies. FHAR1 antibodies are
CC useful to isolate and to identify clones expressing the polypeptides, or
CC to purify the polypeptides by affinity chromatography and to treat cancer
XX
XX Sequence 664 AA;
SQ
Query Match 100.0%; Score 378; DB 4; Length 664;
Best Local Similarity 100.0%; Pred. No. 2.7e-40;
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTEDTSGTVINKLVV 60
DB 31 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIIVVDEKSGQVLTEDTSGTVINKLVV 90
QY 61 KKQTCPLQTGDVI 73
DB 91 KKQTCPLQTGDVI 103
RESULT 7
AAB20219
ID AAB20219 standard; protein; 664 AA.
XX
AC AAB20219;
XX
XX 14-MAY-2001 (first entry)
DT
XX Human Chfr (checkpoint with FHA and ring finger) protein.
DE
XX Checkpoint with forkhead associated domain and ring finger; Chfr; human;
KW mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening;
XX ubiquitin-protein ligase.
XX

```
OS XX Homo sapiens.
FH Key Location/Qualifiers
FT Domain 31..103
FT /label= Forkhead-associated_domain
FT Domain 303..346
FT /label= Ring_finger-domain
FT Region 476..641
FT /note= "cysteine-rich region"
FT Misc-difference 580
FT /note= "Met in U2OS cells"
XX
XX WO200109150-A2.
XX
XX 08-FEB-2001.
XX
XX 14-JUN-2000; 2000WO-US016391.
XX
XX 29-JUL-1999; 99US-0146194P.
XX
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX Halazonetis T, Scolnick D;
XX
XX WPI; 2001-182927/18.
XX N-PSDB; AAF30352.
XX
XX Novel nucleic acid sequence of mitotic checkpoint gene encoding a
XX checkpoint with forkhead-associated domain and ring finger protein, for
XX diagnosing tumorigenic cells and in screening for anticancer drugs.
XX
XX Claim 8(a); Fig 4A-C; 85pp; English.
XX
XX The present sequence is that of human mitotic checkpoint protein Chfr,
XX having a forkhead associated domain (PHA) and a ring finger domain. The
XX protein is required for regulation of the transition of cells from
XX prophase to metaphase during mitosis. It has ubiquitin-protein ligase
XX activity. The Chfr checkpoint was evident in primary human cells, but was
XX inactivated in 4 of 8 human cancer cell lines. In U2OS cells, a mutation
XX was identified that caused a Val to Met amino acid substitution in the
XX highly conserved C-terminal Cys-rich region of the Chfr protein. In the
XX absence of the Chfr checkpoint, cells subjected to mitotic stress
XX condensed their chromosomes despite failing to separate their
XX chromosomes. Chfr may monitor centrosome separation. Inactivation of the
XX Chfr gene (see AAF30352) in human cancer is theorized to underlie the
XX increased sensitivity of cancer cells to antimitotic drugs. Polypeptides
XX comprising the present sequence, or sequences comprising at least amino
XX acids 31-103, 303-346 and/or 476-641 of this sequence, are claimed.
XX Claimed methods of determining the tumorigenic potential of a cell
XX comprise examining the cell for the presence of Chfr expression or for
XX the presence of Chfr-mediated ubiquitin-protein ligase activity (in both
XX cases, absence of expression indicating predisposition to tumourigenesis
XX upon exposure to mitotic stress). A diagnostic kit for detecting the
XX tumorigenic potential of cell cells comprises may comprise a ligand that
XX binds to Chfr, such as an antibody or its fragment. Inhibitors of Chfr
XX are identified by monitoring their effect on Chfr expression, and are
XX used to retard the growth of cancer cells
XX
XX Sequence 664 AA;
XX
XX Query Match 100.0%; Score 378; DB 4; Length 664;
XX Best Local Similarity 100.0%; Pred No. 2,7e-40;
XX Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 VLLRKEWTIGRRRGCDLFFPSNKLVSQDHCRIIVDEKSGQVLTSTSGTVINKLKV 60
XX 31 VLLRKEWTIGRRRGCDLFFPSNKLVSQDHCRIIVDEKSGQVLTSTSGTVINKLKV 90
XX
XX 61 KKQTCPLQGDVI 73
XX 91 KKQTCPLQGDVI 103
XX
XX
```

PR	08-SEP-2000;	2000US-02312433P;
PR	08-SEP-2000;	2000US-02312443P;
PR	08-SEP-2000;	2000US-0231413P;
PR	08-SEP-2000;	2000US-0231414P;
PR	08-SEP-2000;	2000US-0232080P;
PR	08-SEP-2000;	2000US-0232081P;
PR	12-SEP-2000;	2000US-0231968P;
PR	14-SEP-2000;	2000US-0233397P;
PR	14-SEP-2000;	2000US-0233398P;
PR	14-SEP-2000;	2000US-0233399P;
PR	14-SEP-2000;	2000US-0234401P;
PR	14-SEP-2000;	2000US-0234406P;
PR	14-SEP-2000;	2000US-0233063P;
PR	14-SEP-2000;	2000US-0233065P;
PR	21-SEP-2000;	2000US-0234223P;
PR	21-SEP-2000;	2000US-0234223P;
PR	21-SEP-2000;	2000US-0234224P;
PR	25-SEP-2000;	2000US-0234998P;
PR	25-SEP-2000;	2000US-0234997P;
PR	25-SEP-2000;	2000US-0235484P;
PR	27-SEP-2000;	2000US-0235834P;
PR	27-SEP-2000;	2000US-0235836P;
PR	29-SEP-2000;	2000US-0236327P;
PR	29-SEP-2000;	2000US-0236367P;
PR	29-SEP-2000;	2000US-0236368P;
PR	29-SEP-2000;	2000US-0236369P;
PR	29-SEP-2000;	2000US-0236370P;
PR	02-OCT-2000;	2000US-0236802P;
PR	02-OCT-2000;	2000US-0237037P;
PR	02-OCT-2000;	2000US-0237038P;
PR	02-OCT-2000;	2000US-0237039P;
PR	02-OCT-2000;	2000US-0237040P;
PR	13-OCT-2000;	2000US-0239335P;
PR	13-OCT-2000;	2000US-0239337P;
PR	20-OCT-2000;	2000US-0240960P;
PR	20-OCT-2000;	2000US-0241221P;
PR	20-OCT-2000;	2000US-0241785P;
PR	20-OCT-2000;	2000US-0241786P;
PR	20-OCT-2000;	2000US-0241787P;
PR	20-OCT-2000;	2000US-0241808P;
PR	20-OCT-2000;	2000US-0241809P;
PR	01-NOV-2000;	2000US-0244182P;
PR	08-NOV-2000;	2000US-0246474P;
PR	08-NOV-2000;	2000US-0246475P;
PR	08-NOV-2000;	2000US-0246476P;
PR	08-NOV-2000;	2000US-0246477P;
PR	08-NOV-2000;	2000US-0246478P;
PR	08-NOV-2000;	2000US-0246523P;
PR	08-NOV-2000;	2000US-0246524P;
PR	08-NOV-2000;	2000US-0246525P;
PR	08-NOV-2000;	2000US-0246526P;
PR	08-NOV-2000;	2000US-0246527P;
PR	08-NOV-2000;	2000US-0246528P;
PR	08-NOV-2000;	2000US-0246532P;
PR	08-NOV-2000;	2000US-0246609P;
PR	08-NOV-2000;	2000US-0246610P;
PR	08-NOV-2000;	2000US-0246611P;
PR	08-NOV-2000;	2000US-0246613P;
PR	17-NOV-2000;	2000US-0249207P;
PR	17-NOV-2000;	2000US-0249208P;
PR	17-NOV-2000;	2000US-0249209P;
PR	17-NOV-2000;	2000US-0249210P;
PR	17-NOV-2000;	2000US-0249211P;
PR	17-NOV-2000;	2000US-0249212P;
PR	17-NOV-2000;	2000US-0249213P;
PR	17-NOV-2000;	2000US-0249214P;
PR	17-NOV-2000;	2000US-0249215P;
PR	17-NOV-2000;	2000US-0249216P;
PR	17-NOV-2000;	2000US-0249217P;
PR	17-NOV-2000;	2000US-0249218P;
PR	17-NOV-2000;	2000US-0249244P;
PR	17-NOV-2000;	2000US-0249245P;

PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX N-PSDB; AAS26143.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1109; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cardiac
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence data for this patent did not form part of the invention. Note: The
CC sequence data for this patent did not form part of the printed

```

Query Match      80.4%; Score 304; DB 4; Length 92;
Best Local Similarity 98.3%; Pred. No. 1.1e-31;
Matches 59: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Qy	V	L	L	K	R	E	W	T	I	G	R	R	G	C	D	L	S	P	S	N	K	L	V	S	G	D	H	C	R	I	V	D	E	K	S	G	Q	V	T	L	E	D	T	S	T	S	G	T	V	I	N	K	L	K	V					
Db	V	L	L	K	R	E	W	T	I	G	R	R	G	C	D	L	S	P	S	N	K	L	V	S	G	D	H	C	R	I	V	D	E	K	S	G	Q	V	T	L	E	D	T	S	T	S	G	T	V	I	N	K	L	K	V					

RESULT 9	
ABU55225	
ID	ABU55225 standard; protein; 92 AA.
XX	
AC	ABU55225; 92
XX	
DT	18-MAR-2003 (first entry)

XX DE Human novel polypeptide #312.
 XX KW Human; neural disorder; immune system disorder; renal disorder;
 XX KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 XX OS Homo sapiens.
 XX PN US2002132753-A1.
 XX PD 19-SEP-2002.
 XX PF 17-JAN-2001; 2001US-00764864.
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0215647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225477P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 28-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239355P.
 PR 20-OCT-2000; 2000US-0240360P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.

XX PI Rosen CA, Ruben SM, Barash SC;
 XX DR WPI; 2003-147444/14.
 XX DR N-PSDB; ABX73484.
 XX New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX Claim 11; SEQ ID NO 1109; 402pp; English.
 XX The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention
 XX Sequence 92 AA;
 SQ
 Query Match 80.4%; Score 304; DB 6; Length 92;
 Best Local Similarity 98.3%; Pred. No. 1.1e-31;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIWVDEKSGQVLTETSTSGTINKLVV 60
 DB 33 VLLKREWTIGRRGCDLSPFNKLVSGDHCRIWVDEKSGQVLTETSTSGTINKLVV 92
 RESULT 10
 AAU16157
 ID AAU16157 standard; protein; 99 AA.
 AC AAU16157;
 XX 07-NOV-2001 (first entry)
 DT Human novel secreted protein, Seq ID 1110.
 DE Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
 KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
 KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
 KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 KW cerebral ischaemia; angiogenesis; nervous system disorder;
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;
 KW wound healing; epithelial cell proliferation; skin ageing; food additive;
 KW preservative; antiproliferative.
 XX Homo sapiens.
 OS WO200155322-A2.
 XX 02-AUG-2001.
 PD 17-JAN-2001; 2001WO-US001341.
 PF 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0211486P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-02311242P.
PR 08-SEP-2000; 2000US-02311243P.
PR 08-SEP-2000; 2000US-02311244P.
PR 08-SEP-2000; 2000US-02311413P.
PR 08-SEP-2000; 2000US-02311414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-02311968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.
N-PSDB; AAS26144.

New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives.

Claim 11; SEQ ID NO 1110; 980pp; English.

The invention relates to isolated nucleic acid molecules and their

CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 72.8%; Score 275; DB 4; Length 99;
Best Local Similarity 90.3%; Pred. No. 7.3e-28;
Matches 56; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
QY 12 RRRGDLSPFNKLVSGDHCRIVVDKSGQVLTSTSGTVINKLVKKQTCPLQTGD 71
DB 9 RDKGPD--FPXNKLVSQDHCRIVVDKSGQVLTSTSGTVINKLVKKQTCPLQTGD 66
QY 72 VI 73
DB 67 VI 68

RESULT 11
AAU16576
ID AAU16576 standard; protein; 99 AA.

XX AC AAU16576;
XX DT 07-NOV-2001 (first entry)
XX DE Human novel secreted protein, Seq ID 1529.
XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cerebrovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX OS Homo sapiens.
XX PN WO200153322-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US0001341.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184654P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-021680P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
20-OCT-2000; 2000US-0241826P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249298P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250161P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254037P.
05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI, 2001-483783/53.
N-PSDB; AAS25563.
New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives.
Claim 11; SEQ ID NO 1529; 980pp; English.
The invention relates to isolated nucleic acid molecules and their
encoded secreted proteins. The nucleic acids and proteins are used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, cats, dogs, chickens or sheep. They are also used

CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infectious caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Query Match 72.8%; Score 275; DB 4; Length 99;
Best Local Similarity 90.3%; Pred. No. 7,3e-28;
Matches 56; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
QY 12 RRRGCDLSPPSNKLVSGBDCHRIIVVDEKSGQVLTSTSGTVINKLVKKQTCPLQTGD 71
Db 9 RDXGPD--PPXNKLVSGBDCHRIIVVDEKSGQVLTSTSGTVINKLVKKQTCPLQTGD 66
QY 72 VI 73
Db 67 VI 68
RESULT 12
ABUS55645
ID ABUS55645 standard; protein; 99 AA.
XX ABUS55645;
XX 18-MAR-2003 (first entry)
XX Human novel polypeptide #732.
XX Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX Homo sapiens.
XX US2002132753-A1.
XX 19-SEP-2002.
XX 17-JAN-2001; 2001US-00764864.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR N-PSDB; ABX73904.
XX
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX Claim 11; SEQ ID NO 1529; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
XX (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
XX leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
XX appendicitis), allergic reactions and conditions (e.g. asthma), blood
XX related disorders (e.g. thrombosis, atherosclerosis and myocardial
XX infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
XX ABU55748 represent human novel polypeptides of the invention
XX
XX Sequence 99 AA;
SQ

Query Match 72.8%; Score 275; DB 6; Length 99;

Best Local Similarity 90.3%; Pred. No. 7.3e-28;
Matches 56; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
QY 12 RREGCDLSFSPSKLVSGDHCRIIVVDKSGQVLTEDTSTGTVINKLKVKKQTCPLQTGD 71
DB 9 RDKGPD--FFXKNLVSGDHCRIIVVDKSGQVLTEDTSTGTVINKLKVKKQTCPLQTGD 66
QY 72 VI 73
DB 67 VI 68
RESULT 13
ABU55226
ID ABU55226 standard; protein; 99 AA.
XX
XX AC ABU55226;
XX
XX DT 18-MAR-2003 (first entry)
XX
XX DE Human novel polypeptide #313.
XX
XX Human; neural disorder; immune system disorder; renal disorder;
XX muscular disorder; respiratory disease; reproductive disorder;
XX gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
XX hyperproliferative disorder; inflammatory disease; allergic reaction;
XX blood related disorder; cancer; immunosuppressive; antiinflammatory;
XX cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
XX haemostatic; antiarteriosclerotic.
XX
XX OS Homo sapiens.
XX
XX PN US2002132753-A1.
XX
XX PD 19-SEP-2002.
XX
XX PF 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-147444/14.
 DR N-PSDB; ABX73485.
 XX
 PT New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX
 PS Claim 11; SEQ ID NO 1110; 402pp; English.
 CC
 CC The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention
 XX
 SQ Sequence 99 AA;
 Query Match 72.8%; Score 275; DB 6; Length 99;
 Best Local Similarity 90.3%; Pred. No. 7.3e-28;
 Matches 56; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
 QY 12 RRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVKVKQTCPLQTGD 71
 Db 9 RDXGPD--FPFNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVKVKQTCPLQTGD 66
 QY 72 VI 73
 Db 67 VI 68
 RESULT 14
 AAW73894
 ID AAW73894 standard; protein; 776 AA.
 XX
 AC AAW73894;
 XX
 DT 08-APR-1999 (first entry)
 XX
 DE Yeast MEC2 protein sequence.
 XX

KW Cell cycle checkpoint gene; yeast; radiation resistance; detection;
 KW G2/M checkpoint.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN US5866338-A.
 XX
 PD 02-FEB-1999.
 XX
 PF 06-JUN-1997; 97US-00870693.
 XX
 PR 12-MAY-1992; 92US-00882051.
 PR 14-MAY-1992; 92US-00884426.
 PR 12-MAY-1993; 93WO-US004458.
 PR 18-FEB-1994; 94US-00198446.
 XX
 PA (UNIW) UNIV WASHINGTON.
 PA (UYAR-) UNIV ARIZONA.
 PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
 XX
 PI Plon SE, Groudine MT, Hartwell LH, Weinert TA;
 XX
 DR WPI; 1999-141919/12.
 DR N-PSDB; AAX01275.
 XX
 PT Nucleotide sequences that hybridise to huRAD cDNA - method for isolating
 PT human checkpoint cDNA, and antibody to CDC34.
 XX
 PS Example 4; Col 89-96; 73pp; English.
 XX
 CC This sequence is the yeast MEC2 protein. The DNA encoding this sequence
 CC is necessary for G2/M cell cycle checkpoint control. This sequence was
 CC identified using the method of the invention for isolating a human
 CC checkpoint cDNA that is capable of restoring growth at a restrictive
 CC temperature in a yeast test cell, where the yeast test cell comprises a
 CC genome having a first gene that forms a DNA strand break at a restrictive
 CC temperature and a second gene that fails to induce a cell cycle arrest in
 CC response to the DNA strand break, whereby the growth of the yeast test
 CC cell is inhibited at the restrictive temperature, comprising: (a)
 CC obtaining a human cDNA library comprising several human cDNA clones; (b)
 CC inserting the human cDNA clones individually into plasmid vectors
 CC containing a selectable marker gene; (c) transforming a culture of the
 CC yeast test cells with the plasmid vectors from the preceding step; (d)
 CC selecting for yeast test cells transformed with the selectable marker
 CC gene; (e) growing the selected transformants at the restrictive
 CC temperature and isolating a candidate transformant capable of growing at
 CC the restrictive temperature; and (f) identifying the human cDNA carried
 CC by the candidate transformant as a human checkpoint cDNA by sequencing that
 CC the human cDNA carried by the candidate transformant and determining that
 CC the human cDNA is less than 50% homologous with both the first gene and
 CC the second gene. The nucleotide sequence of huRADcompS can be used to
 CC confer radiation resistance on a cell
 XX
 SQ Sequence 776 AA;
 Query Match 28.4%; Score 107.5; DB 2; Length 776;
 Best Local Similarity 37.7%; Pred. No. 8.7e-05;
 Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;
 QY 5 KREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVKVKQTC 64
 Db 63 KKWTFGPNPACDYLHGNISRLSNKHFOILLGE-DGNLLINDISTNGTWNQGVKNSN 121
 QY 65 CPLQTGDVI 73
 Db 122 QLLSQGDEI 130
 RESULT 15
 AAW26664
 ID AAW26664 standard; protein; 821 AA.
 XX
 AC AAW26664;

Search completed: May 7, 2004, 14:45:46
Job time : 17.2046 secs

XX 25-MAR-2003 (revised)

DT 25-FEB-1998 (first entry)

XX

DE Yeast checkpoint control protein MEC2.

XX

XX MEC2; cell cycle; checkpoint gene; yeast; DNA damage; cancer; therapy.

XX

XX Saccharomyces cerevisiae.

XX

XX Location/Qualifiers

FT 1..776

FT /note= "in the Sequence Listing MEC2 terminates at

FT residue 775"

XX

XX US5674996-A.

XX

XX 07-OCT-1997.

XX

XX 18-FEB-1994; 94US-00198446.

XX

XX 12-MAY-1992; 92US-00882051.

XX

XX 14-MAY-1992; 92US-00884426.

XX

XX 12-MAY-1993; 93WO-US004456.

XX

XX (UYAR-) UNIV ARIZONA.

XX

XX (UNIW) UNIV WASHINGTON.

XX

XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.

XX

XX Groudine MT, Weinert TA, Plon SE, Hartwell LH;

XX

XX WPI; 1997-502392/46.

XX

XX N-PSDB; AAT91040.

XX

XX Nucleotide sequence capable of hybridising with huCDC34 - which is human

XX

XX checkpoint gene, useful to increase sensitivity of tumour cells to

XX

XX chemotherapeutic drugs or radiation.

XX

XX Example 4; Col 85-90; 54pp; English.

XX

XX This protein is encoded by novel yeast checkpoint control gene MEC2 (see

XX

XX AAT91040). Yeast checkpoint control genes RAD17, RAD24, MEC1, MEC2 and

XX

XX MEC3 (AAT91037-41, respectively) are responsible for recognising if the

XX

XX cell has suffered DNA damage in the form of radiation or chemical damage

XX

XX or if the cell has failed to complete DNA replication because of chemical

XX

XX inhibition or intrinsic error. Upon recognising damage or failure, the

XX

XX genes are responsible for inhibiting mitosis. The purpose of this

XX

XX checkpoint control is that it preserves the viability of the cell and the

XX

XX integrity of the genome by providing the cell time to repair these

XX

XX insults prior to undertaking mitosis. The genes are potentially useful in

XX

XX developing cancer chemotherapeutics, cancer chemoprevention agents, and

XX

XX environmental toxicology tests. They can be used to produce proteins (see

XX

XX AAW26661-65) that can then be screened for chemical agents that would

XX

XX interfere with checkpoint controls. Cloned genes can also be used to

XX

XX develop yeast strains in which these genes are deleted. Such yeast

XX

XX strains can then be used to find the homologous human genes (see AAT91034

XX

XX -36). (Updated on 25-MAR-2003 to correct PF field.)

XX

XX Sequence 821 AA;

XX

XX Query Match 28.4%; Score 107.5; DB 2; Length 821;

XX

XX Best Local Similarity 37.7%; Pred. No. 9.3e-05;

XX

XX Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;

XX

XX 5 KREWTIGRRGCDLSPFSKNLVSCHRCIVVDEKSGOVTLDTSTGTINKLVKKQT 64

XX

XX 63 KKVWTFGRNACDYHLGNISRLSNKHFQILLGE-DGNILLNDISTNGTWNQKRVKNSN 121

XX

XX 65 CPLQTGVPI 73

XX

XX 122 QLLSQGDEI 130

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:42:23 ; Search time 4.85639 Seconds
(without alignments)
776.028 Million cell updates/sec

Title: US-10-048-046-2_COPY_31_103

Perfect score: 378
Sequence: 1 VLLKRGWTIGRRGCDLSF.....INKLKVVVKQTCPLQTGQDVI 73

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgm2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgm2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgm2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgm2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgm2_6/ptodata/2/iaa/6C COMB.pcp.*
6: /cgm2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107.5	28.4	776	1	US-08-198-446B-17
2	107.5	28.4	776	2	US-08-870-693-17
3	107.5	28.4	821	1	US-08-198-446B-6
4	107.5	28.4	821	2	US-08-870-693-6
5	97	25.7	65	4	US-09-300-008B-46
6	79.5	21.0	558	4	US-09-252-991A-17202
7	76.5	20.2	517	4	US-09-849-617-2
8	74.5	19.7	64	4	US-09-300-008B-44
9	70.5	18.7	877	4	US-09-328-352-8162
10	68	18.0	754	4	US-09-300-008B-2
11	67	17.7	63	4	US-09-300-008B-43
12	64.5	17.1	64	4	US-09-300-008B-41
13	61	16.1	212	4	US-09-328-352-5652
14	61	16.1	1070	4	US-09-107-532A-6735
15	59.5	15.7	804	3	US-08-855-910-4
16	59.5	15.7	824	4	US-09-134-000C-4908
17	59	15.6	1045	1	US-07-596-467-6
18	59	15.6	1045	1	US-07-934-374-6
19	59	15.6	1045	1	US-07-783-861C-6
20	58.5	15.5	475	4	US-09-252-991A-22693
21	58.5	15.5	502	4	US-09-232-991A-30007
22	58	15.3	721	4	US-08-851-435-2
23	58	15.3	746	4	US-08-851-435-6
24	57.5	15.2	80	4	US-09-540-236-2410
25	57.5	15.2	443	4	US-09-134-001C-3148
26	57.5	15.2	632	3	US-09-354-129-8
27	57.5	15.2	632	4	US-09-504-357-8

28	57	15.1	368	3	US-08-961-083-76	Sequence 76, Appl
29	57	15.1	368	4	US-09-536-784-76	Sequence 76, Appl
30	57	15.1	870	4	US-09-543-681A-5184	Sequence 5184, Ap
31	56.5	14.3	440	4	US-09-198-452A-755	Sequence 755, App
32	56	14.8	543	4	US-09-529-093A-2	Sequence 2, Appl
33	56	14.8	543	4	US-09-529-154-2	Sequence 20, Appl
34	56	14.8	830	3	US-08-804-439A-20	Sequence 20, Appl
35	56	14.8	830	3	US-08-720-229-20	Sequence 13174, A
36	56	14.8	878	4	US-09-489-039A-13174	Sequence 20, Appl
37	56	14.8	913	4	US-09-252-991A-22127	Sequence 22127, A
38	56	14.8	3256	4	US-09-919-172-98	Sequence 98, Appl
39	56	14.8	3256	4	US-09-978-594-22	Sequence 22, Appl
40	55.5	14.7	1548	4	US-09-252-991A-22301	Sequence 22301, A
41	55	14.6	65	4	US-09-300-008B-48	Sequence 48, Appl
42	55	14.6	283	4	US-09-107-532A-6481	Sequence 6481, Ap
43	55	14.6	312	4	US-09-107-532A-5790	Sequence 5790, Ap
44	55	14.6	356	4	US-09-540-236-3075	Sequence 3075, Ap
45	55	14.6	1302	4	US-09-423-890-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-198-446B-17
; Sequence 17, Application US/08198446B
; Patent No. 5674996
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/198.446B
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHCRI7537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 776 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: Yeast MEC2 protein
US-08-198-446B-17

Query Match 28.4%; Score 107.5; DB 1; Length 776;
Best Local Similarity 37.7%; Pred. No. 2.4e-05;
Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;
OY 5 KREWTIGRRGCDLSFPSPNKLVSVDKSGQVLTEDTSGTVINKLVKVKQT 64
Db 63 KKVTFGRNPACDYLHGNISRLSNKHFOILLGE-DGNLLINDISTNGTWLNGQVKVKN 121

QY 65 CPLQTDGVI 73
DB 122 QLLSQGDEI 130

RESULT 2

US-08-870-693-17
; Sequence 17, Application US/08870693
; Patent No. 5866338
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/870,693
FILING DATE:
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/198,446
FILING DATE: February 18, 1994
APPLICATION NUMBER: PCT/US93/04458
FILING DATE: May 12, 1993
APPLICATION NUMBER: US 07/884,426
FILING DATE: May 14, 1992
APPLICATION NUMBER: US 07/882,051
FILING DATE: May 12, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCRI10798

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-224-0779
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Yeast MEC2 protein

US-08-870-693-17

Query Match 28.4%; Score 107.5; DB 2; Length 776;
Best Local Similarity 37.7%; Pred. No. 2.4e-05;
Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;

QY 5 KREWTIGRRGCDLSFFPSNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVKKQT 64
DB 63 KKWTFGRNPACDYLGNISRLSNKHFQILLGE-DGNLLNDISTNGTWNLGQVKNSN 131
QY 65 CPLQTDGVI 73
DB 122 QLLSQGDEI 130

RESULT 3

US-08-198-446B-6

; Sequence 6, Application US/08198446B
; Patent No. 5674996
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/198,446B
FILING DATE: 18-FEB-1994
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCRI17537
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-224-0779

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 821 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Yeast MEC2 protein

US-08-198-446B-6

Query Match 28.4%; Score 107.5; DB 1; Length 821;
Best Local Similarity 37.7%; Pred. No. 2.6e-05;
Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;

QY 5 KREWTIGRRGCDLSFFPSNKLVSGDHCRIVVDEKSGQVLTEDTSTSGTVINKLVKKQT 64
DB 63 KKWTFGRNPACDYLGNISRLSNKHFQILLGE-DGNLLNDISTNGTWNLGQVKNSN 121
QY 65 CPLQTDGVI 73
DB 122 QLLSQGDEI 130

RESULT 4

US-08-870-693-6
; Sequence 6, Application US/08870693

; Patent No. 5866338
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Plon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/870,693
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
FILING DATE: February 18, 1994
APPLICATION NUMBER: PCT/US93/04458
FILING DATE: May 12, 1993
APPLICATION NUMBER: US 07/884,426
FILING DATE: May 14, 1992
APPLICATION NUMBER: US 07/882,051
FILING DATE: May 12, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCRI10798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-224-0779
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 821 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Yeast MEC2 protein
US-08-870-693-6

Query Match 28.4%; Score 107.5; DB 2; Length 821;
Best Local Similarity 37.7%; Pred. No. 2.6e-05;
Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;

Qy 5 KREWTIGRRGCDLSFSPNKLVS GDHCRIVVDEKSGQVLTEDTSTGTVINKLVKVKQT 64
Db 63 KKVWTFGRNPACDYHLGNISRLSNKHFQILLGE-DGNLLNDISTNGTWLNGQVKVNSN 121

Qy 55 CPLQTGDVI 73
Db 122 QLLSQGDEI 130

RESULT 5
US-09-300-008B-46
Sequence 46, Application US/09300008B
Patent No. 6458334
GENERAL INFORMATION:
APPLICANT: Concannon et al.
TITLE OF INVENTION: A GENE ASSOCIATED WITH NUMJEN BREAKAGE
TITLE OF INVENTION: SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE
FILE REFERENCE: 9924-0003-228
CURRENT APPLICATION NUMBER: US/09/300,008B
CURRENT FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: US 60/083,269
PRIOR FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 46
LENGTH: 65
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-300-008B-46

Query Match 25.7%; Score 97; DB 4; Length 65;
Best Local Similarity 36.4%; Pred. No. 2.3e-05;
Matches 24; Conservative 8; Mismatches 30; Indels 4; Gaps 1;

Qy 8 WTIGRRGCDLSFSPNKLVS GDHCRIVVDEKSGQVLTEDTSTGTVINKLVKVKQTCP 67
Db 1 WTFGRNPACDYHLGNISRLSNKHFQILL-----GNLLNDISTNGTWLNGQVKVNSNOLL 56

Qy 68 QTGDVI 73
Db 57 SQGDEI 62

RESULT 6
US-09-252-991A-17202
Sequence 17202, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17202
LENGTH: 558
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17202

Query Match 21.0%; Score 79.5; DB 4; Length 558;
Best Local Similarity 32.5%; Pred. No. 0.07;
Matches 25; Conservative 16; Mismatches 27; Indels 9; Gaps 4;

Qy 2 LURKREWTIGRRGCDLSFSPNKLVS GDHCRIVVDEKSGQVLTEDTSTGTVI-----NK 56
Db 85 VIDQGLTIGRGFDNDWVLPDPERLVSSRHCTIL--NRDGVYLTDTSTNGVLLVNAHR 142

Qy 57 LKVVKKQTCPLQTGDVI 73
Db 143 LR--RGNSEPLQDGETV 157

RESULT 7
US-09-849-617-2
Sequence 2, Application US/09849617
Patent No. 6593110
GENERAL INFORMATION:
APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
APPLICANT: GUO, Zijian
APPLICANT: DUNPHY, William
TITLE OF INVENTION: CHECKPOINT-ACTIVATING OLIGONUCLEOTIDES
FILE REFERENCE: CITI1350-1
CURRENT APPLICATION NUMBER: US/09/849,617
CURRENT FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: US 60/202,028
PRIOR FILING DATE: 2000-05-04
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 517
TYPE: PRT
ORGANISM: Xenopus laevis
US-09-849-617-2

Query Match 20.2%; Score 76.5; DB 4; Length 517;
Best Local Similarity 31.9%; Pred. No. 0.16;
Matches 23; Conservative 9; Mismatches 29; Indels 11; Gaps 2;

Qy 7 EWTIGRRGCDLSF-----PSNKLVS GDHCRIVVDEKSGQ---VTLEDTSTGTVIN 55
Db 84 EYVFGDKCDYTFDIPVLNQTDRYKTYKRFHFQELGHSRVRANIEDLSGNGTFVN 143

```
QY 56 KLKVKVKQTCTPL 67
Db 144 KBIIGKRTPL 155

RESULT 8
US-09-300-008B-44
; Sequence 44, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-09-300-008B-44

Query Match 19.7%; Score 74.5; DB 4; Length 64;
Best Local Similarity 31.8%; Pred. No. 0.019;
Matches 21; Conservative 9; Mismatches 31; Indels 5; Gaps 2;

QY 8 WTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTSTSGTVINKLVKVKQTCTPL 67
Db 1 WGFGRHKSCVVL-NGPRVSNFHEIY----QGVVFLHDHSSNGTFLNFERLAKNSRTIL 55

QY 68 QTGDVI 73
Db 56 SNGDEI 61

RESULT 9
US-09-328-352-8162
; Sequence 8162, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8162
; LENGTH: 877
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8162

Query Match 18.7%; Score 70.5; DB 4; Length 877;
Best Local Similarity 25.9%; Pred. No. 1.9;
Matches 22; Conservative 12; Mismatches 24; Indels 27; Gaps 4;

QY 1 VLLRKREWTIGRR--GCDLSPFNKLVSGDHCRIVVDEKSGQVLTSTST-----49
Db 430 VQFLRDWGVSRQRYWGCPI-----PMINCDTC-----GQVTPSDQLPVLPDVI 476

QY 50 ---SGTVINKLVKVKQTCTPLQTD 71
Db 477 PDGSGNPLKNKFEFFETKPCCGD 501

RESULT 10
US-09-300-008B-2
; Sequence 2, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-300-008B-2

Query Match 18.0%; Score 68; DB 4; Length 754;
Best Local Similarity 24.7%; Pred. No. 3.2;
Matches 20; Conservative 21; Mismatches 30; Indels 10; Gaps 2;

QY 2 LLRKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTSTSTSGT 52
Db 18 LLTGVEYVVG-RKNCALLIENDQISGRNHAULTANFSVTNLSTQDEIPVLTLDKNSKYGT 76

QY 53 VINKLVKVKQTCTPLQTDVI 73
Db 77 FVNEERMQNGFSRTLKSGDGI 97

RESULT 11
US-09-300-008B-43
; Sequence 43, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-300-008B-43

Query Match 17.7%; Score 67; DB 4; Length 63;
Best Local Similarity 34.3%; Pred. No. 0.18;
Matches 23; Conservative 10; Mismatches 24; Indels 10; Gaps 3;

QY 9 TIGERRGCD--LSFPFNKLVSGDHCRIVVDEKSGQVLTSTSGTVINKLVKVKQTCTP 66
Db 2 TIGRSCDVLSEPDISTFHFHLLIN-----VIDKSRNGTFINGRLVKQDI- 53

QY 67 LOTGDVI 73
Db 54 LKNGDRI 60

RESULT 12
US-09-300-008B-41
; Sequence 41, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
```

```
; Sequence 2, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-300-008B-2

Query Match 18.0%; Score 68; DB 4; Length 754;
Best Local Similarity 24.7%; Pred. No. 3.2;
Matches 20; Conservative 21; Mismatches 30; Indels 10; Gaps 2;

QY 2 LLRKREWTIGRRGCDLSPFNKLVSGDHCRIVVDEKSGQVLTSTSTSGT 52
Db 18 LLTGVEYVVG-RKNCALLIENDQISGRNHAULTANFSVTNLSTQDEIPVLTLDKNSKYGT 76

QY 53 VINKLVKVKQTCTPLQTDVI 73
Db 77 FVNEERMQNGFSRTLKSGDGI 97

RESULT 11
US-09-300-008B-43
; Sequence 43, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-300-008B-43

Query Match 17.7%; Score 67; DB 4; Length 63;
Best Local Similarity 34.3%; Pred. No. 0.18;
Matches 23; Conservative 10; Mismatches 24; Indels 10; Gaps 3;

QY 9 TIGERRGCD--LSFPFNKLVSGDHCRIVVDEKSGQVLTSTSGTVINKLVKVKQTCTP 66
Db 2 TIGRSCDVLSEPDISTFHFHLLIN-----VIDKSRNGTFINGRLVKQDI- 53

QY 67 LOTGDVI 73
Db 54 LKNGDRI 60

RESULT 12
US-09-300-008B-41
; Sequence 41, Application US/09300008B
; Patent No. 6458534
; GENERAL INFORMATION:
; APPLICANT: Concannon et al.
; TITLE OF INVENTION: A GENE ASSOCIATED WITH NIJMEGEN BREAKAGE
```

; TITLE OF INVENTION: SYNDROME, ITS GENE PRODUCT AND METHODS FOR THEIR USE
; FILE REFERENCE: 9924-0003-228
; CURRENT APPLICATION NUMBER: US/09/300,008B
; CURRENT FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: US 60/083,269
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-300-008B-41

Query Match 17.1%; Score 64.5; DB 4; Length 64;
Best Local Similarity 25.8%; Pred. No. 0.38; 25; Indels 5; Gaps 2;
Matches 17; Conservative 19; Mismatches 25; Indels 5; Gaps 2;

QY 8 WTIGRRGCDLSPFNKLVGSDHCRIVVDEKSGQVLTSTGTIVINKLVKVKQTCPL 67
DB 1 YVVG-RKNCALLIENDQSISRHAULT---ANVLTKDNSKYGTTFVNEEKMGNGFSRTL 55

QY 68 QTGDVI 73
DB 56 KSGDGI 61

RESULT 13
US-09-328-352-5652
; Sequence 5652, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5652
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-5652

Query Match 16.1%; Score 61; DB 4; Length 212;
Best Local Similarity 23.2%; Pred. No. 5.1;
Matches 16; Conservative 24; Mismatches 23; Indels 6; Gaps 4;

QY 6 REWTIGRRGCDLSPFNKLVGSDHCRIVVDEKSGQVLTED-TSTSGTVINKLVKVKQTC 64
DB 24 RDLVGRHQDADLLQAAE-ISRHAALLKQQA--LWVQDLNLSNGTFFVNDMRI--EQE 78

QY 65 CPLQTGDVI 73
DB 79 KQLHGDIV 87

RESULT 14
US-09-107-532A-6735
; Sequence 6735, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA

; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6735:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1070 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...1070
; SEQUENCE DESCRIPTION: SEQ ID NO: 6735:
US-09-107-532A-6735

Query Match 16.1%; Score 61; DB 4; Length 1070;
Best Local Similarity 42.3%; Pred. No. 41;
Matches 11; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 18 LSPFNKLVGSDHCRIVVDEKSGQVT 43
DB 866 LSIPEQLFAGDVLIVLNDQAGEAT 891

RESULT 15
US-08-855-910-4
; Sequence 4, Application US/08855910
; Patent No. 6221640
; GENERAL INFORMATION:
; APPLICANT: Tao, Jianshi
; APPLICANT: Sassanfar, Mandana
; APPLICANT: Gallant, Paul L.
; APPLICANT: Shen, Xiaoyu
; APPLICANT: Avruich, Anthony S.
; APPLICANT: Yu, Russell V.
; APPLICANT: Nair, Shamila
; TITLE OF INVENTION: ENTEROCOCCAL AMINOACYL-tRNA SYNTHETASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/855,910
FILING DATE: 14-MAY-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI95-08
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 804 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-855-910-4

Query Match 15.7%; Score 59.5; DB 3; Length 804;
Best Local Similarity 31.5%; Pred. No. 45;
Matches 23; Conservative 12; Mismatches 21; Indels 17; Gaps 4;
QY 5 KREWITGRRGCDLSPFNKLVSGDHCRIWVDEKSGQV--TLEDT---STSGTVINKLV 59
Db 218 QRNW-IGRSEGANVT-----KVAGTESEFTVTRPDTLFGATYTVLAPLELEL 265

QY 60 VKKQTCPLQTGDV 72
Db 266 VKKITTPECTAAV 278

Search completed: May 7, 2004, 14:51:22
Job time : 5.85639 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 14:41:17 ; Search time 4.08553 Seconds
(without alignments)
1718.743 Million cell updates/sec

Title: US-10-048-046-2_COPY_31_103

Perfect score: 378
Sequence: 1 VLLRKEEWIGRRGCDLSF.....INKLKVVKQTCLQTQDVI 73

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: Pir1.*
- 2: Pir2.*
- 3: Pir3.*
- 4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	107.5	28.4	821	1	A39616	protein kinase RAD
2	88	23.3	159	2	A36962	FFA-domain contain
3	79.5	21.0	497	2	F83634	hypothetical prote
4	79	20.9	358	2	A61188	probable transcrip
5	77.5	20.5	460	2	S58882	protein kinase Cds
6	76	20.1	2541	2	T29340	hypothetical prote
7	74.5	19.7	265	2	B61188	Sci protein - mous
8	72	19.0	513	1	S43941	protein kinase DUN
9	71	18.8	409	2	S73738	phosphoglycerate k
10	70.5	18.7	230	2	A93872	hypothetical prote
11	70.5	18.7	301	2	T40402	forkhead nuclear s
12	68.5	18.1	157	2	S76488	hypothetical prote
13	68	18.0	399	2	AC3089	conserved hypothet
14	68	18.0	399	2	G98197	hypothetical prote
15	68	18.0	754	2	T00393	Nijmegen breakage
16	67.5	17.9	234	2	AF2214	hypothetical prote
17	66.5	17.6	546	2	A51899	adenylate cyclase
18	66	17.5	287	2	E69901	cell wall-binding
19	66	17.5	604	2	S52706	probable membrane
20	66	17.5	692	2	T00025	PSP-95 binding pro
21	66	17.5	977	2	T00014	DAP-1 alpha protei
22	64	16.9	1441	1	GNVUSV	M polyprotein prec
23	63.5	16.8	445	2	T43420	probable protein k
24	63.5	16.8	938	2	C84480	hypothetical prote
25	62.5	16.5	357	2	H82645	conserved hypothet
26	62.5	16.5	468	2	D96904	probable membrane
27	62.5	16.5	527	2	B70700	hypothetical prote
28	62	16.4	273	2	AF0041	L-thanosase operon
29	62	16.4	749	2	S75331	penicillin-binding

30	62	16.4	1471	2	D72758	hypothetical prote
31	62	16.4	1637	2	T00070	hypothetical prote
32	62	16.4	2938	2	T30249	cell proliferation
33	61.5	16.3	287	2	H64690	type IIS restricti
34	61.5	16.3	463	2	T10015	hypothetical prote
35	61.5	16.3	488	2	F86911	conserved hypothet
36	61	16.1	484	2	S48403	FKRI protein - yea
37	61	16.1	1433	1	GNVUBW	M polyprotein prec
38	61	16.1	1441	1	GNVULC	M polyprotein prec
39	60.5	16.0	337	2	S75018	adenylate cyclase
40	60.5	16.0	397	2	D83437	hypothetical prote
41	60.5	16.0	661	2	T15073	hypothetical prote
42	60	15.9	353	2	E81680	conserved hypothet
43	60	15.9	560	2	D72478	hypothetical prote
44	60	15.9	952	2	T18837	hypothetical prote
45	60	15.9	969	2	B87083	excinuclease ABC s

ALIGNMENTS

RESULT 1

A39616
protein kinase RAD53 (EC 2.7.1.1-) - Yeast (Saccharomyces cerevisiae)
N/Alternate names: protein P2588; protein YPL153c; SPK1 protein
C/Species: Saccharomyces cerevisiae
C/Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 21-Jul-2000
C/Accession: A39616; S65164; S69446; S13321
R/Stern, D.F.; Zheng, P.; Beidler, D.R.; Zerillo, C.
Mol. Cell. Biol. 11, 987-1001, 1991
A/Title: Spk1, a new kinase from Saccharomyces cerevisiae, phosphorylates proteins on ser
A/Reference number: A39616; MUID:91117267; PMID:1899289
A/Accession: A39616
A/Molecule type: DNA
A/Residues: 1-821 <STE>
A/Cross-references: GB:M55623; NID:gl72656; PIDN:AAA35070.1; PID:gl72657
A/Experimental source: strain S288C
R/Zheng, P.; Fay, D.S.; Burton, J.; Xiao, H.; Pinkham, J.L.; Stern, D.F.
Mol. Cell. Biol. 13, 5829-5842, 1993
A/Title: SPK1 is an essential S-phase-specific gene of Saccharomyces cerevisiae that enc
A/Reference number: A54697; MUID:93361015; PMID:8355715
A/Contents: annotation
R/Purnelle, B.; Coster, F.; Goffeau, A.
submitted to the Protein Sequence Database, May 1996
A/Reference number: S65154
A/Accession: S65164
A/Molecule type: DNA
A/Residues: 1-821 <PUR>
A/Cross-references: EMBL:Z73509; NID:gl370325; PIDN:CAA97858.1; PID:gl370326; GSPDB:GN000
R/Purnelle, B.; Combles, S.; Coster, F.; Naveau, F.; Goffeau, A.
submitted to the EMBL Data Library, March 1996
A/Description: The sequence of 55 kb on the left arm of yeast chromosome XVI identifies ;
ogue to the human phosphotyrosyl phosphatase activator PTPA and a homologue to the plant
A/Reference number: S69446
A/Accession: S69446
A/Molecule type: DNA
A/Residues: 1-821 <PUW>
A/Cross-references: EMBL:X96770; NID:gl403537; PIDN:CAA65568.1; PID:gl403556
C/Genetics:
A/Gene: SGD:RAD53; SPK1; MEC2; SAD1; MIPS:YPL153c
A/Cross-references: SGD:S0006074; MIPS:YPL153c
A/Map position: 16L
C/Function:
A/Description: serine/threonine-specific protein kinase
A/Note: contains low activity as tyrosine-specific protein kinase
C/Superfamily: protein kinase SPK1; kinase interaction domain homology; protein kinase hc
C/Keywords: ATP; cell cycle control; nucleus; phosphotransferase; serine/threonine-specifi
F/68-133/Domain: kinase interaction domain homology <KIH>
F/196-466/Domain: protein kinase homology <KIN>
F/204-213/Region: protein kinase ATP-binding motif
F/227/Active site: Lys #status predicted

Query Match 28.4%; Score 107.5; DB 1; Length 821;
Best Local Similarity 37.7%; Pred.No. 0.00029;
Matches 26; Conservative 9; Mismatches 33; Indels 1; Gaps 1;

QY 5 KREMTIGRRGCDLSFFPSNKLVSGDHCRIVVDEKSGOVTLDETSTGVINKLVKKVKQT 64
| | | | | : | : | : | : | : | : | : | : | : |
Db 63 KKWTTFGRNPACDYHLGNISRLSNKHQPLLGE-DGNLLINDISTNGTILWGKRVKNSN 121

QY 65 CPLQTGGVI 73
| | | | |
Db 122 QLLSQGDEI 130

RESULT 2
A96962
FHA-domain containing secreted protein [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2000 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: A96962
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: A96962
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <CUR>
A:Cross-references: GB:AE001437; PIDN:AAK78484.1; PID:g15023366; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0504

Query Match 23.3%; Score 88; DB 2; Length 159;
Best Local Similarity 41.4%; Pred.No. 0.0086;
Matches 29; Conservative 12; Mismatches 21; Indels 8; Gaps 5;

QY 6 REWITGRRGCDLSFFPSNK-LVSGDHCRIVVDEKSGOVTLLED-TSTSGTVINKLVKKVKQ 63
| | | | | : | : | : | : | : | : | : | : | : |
Db 85 REITIGRKD--DNSIMNEGVSQHARVYL--RNNOYILEDLNSTGTVLNGSKI--KS 138

QY 64 TCPLQTGGVI 73
| | | | |
Db 139 KAVIKSGDEI 148

RESULT 3
F83634
hypothetical protein PA0081 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83634
R:Stover, C.K.; Yuen, Y.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bu
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83634
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-497 <SFO>
A:Cross-references: GB:AE004447; GB:AE004091; NID:g9945902; PIDN:AAG03471.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA0081

Db 22 VLDQGQLTIGRGPDNDWLVLPDPRLVSSRHCTIL--NRDGVYVLTDTSTNGVLLVNAAGR 79

Qy 57 LKVVKEQTCTPLQGTGDVI 73
| : : : : :
Db 80 LR--RGNSEPLQGTGV 94

RESULT 4

A61188
probable transcription factor SC1 - human
C/Species: Homo sapiens (man)
C/Date: 12-May-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1999
C/Accession: A61188
R/Ku, D.H.; Chang, C.; Koniecki, J.; Cannizzaro, L.A.; Boghosian-Sell, L.; Alder, H.; Bae
Cell Growth Differ. 2, 179-186, 1991
A/Title: A new growth-regulated complementary DNA with the sequence of a putative trans-
A/Reference number: A61188; MUID:91329275; PMID:1868030
A/Accession: A61188
A/Molecule type: mRNA
A/Residues: 1-358 <KUA>
A/Note: authors translated the codon TAC for residue 19 as Thr
C/Genetics:
A/Gene: SC1
A/Map position: 6p21-22

Query Match 20.9%; Score 79; DB 2; Length 358;
Best Local Similarity 27.8%; Pred. No. 0.23;
Matches 20; Conservative 20; Mismatches 26; Indels 6; Gaps 4;

Qy 8 WTGRRRG-CDLSF-PSNK--LVSGDHCRIVDEKSG--QVLTEDTSTGTVINKLKVVK 61
| : : : : :
Db 31 YRLGHRADLCVLRPQQEPLGISGTHAEHAEPRGDDWVSLSDHSSQGLVNVRLPR 90
| : : : : :
Qy 62 KQTCPLQGTGDVI 73
| : : : : :
Db 91 GHRLELSDGDL 102

RESULT 5

S58882
protein kinase Cdelta (EC 2.7.1.-) [validated] - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 31-Dec-2000
C/Accession: S58882; S71846; T41204; T52473
R/Murakami, H.; Okayama, H.
Nature 374, 817-819, 1995
A/Title: A kinase from fission yeast responsible for blocking mitosis in S phase.
A/Reference number: S58882; MUID:95240713; PMID:7723827
A/Accession: S58882
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-460 <MURI>
A/Cross-references: EMBL:X85040; NID:G794146
R/Murakami, H.
submitted to the EMBL Data Library, March 1995
A/Reference number: S71846
A/Accession: S71846
A/Molecule type: mRNA
A/Residues: 1-60, 'G', 62-202, 'I', 204-460 <MUR2>
A/Cross-references: EMBL:X85040; NID:G794146; PIDN:CAA59410.1; PID:G794147
R/Wood, V.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.
submitted to the EMBL Data Library, July 1999
A/Reference number: Z21978
A/Accession: T41204
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-202, 'I', 204-237, 'F', 239-460 <WOO>
A/Cross-references: EMBL:AL109736; NID:G5701956; PIDN:CAB52158.1; PID:G5701966; GSPDDB:GN
A/Experimental source: strain 972h(-); cosmid c1855
R/Lindsay, H.D.; Griffiths, D.J.F.; Edwards, R.J.; Christensen, P.U.; Murray, J.M.; Osma
Genes Dev. 12, 382-395, 1998
A/Title: S-phase specific activation of Cdc1 kinase defines a subpathway of the checkpoint
A/Reference number: Z26084; MUID:98119835; PMID:9450932

A;Accession: T52473
A;Status: preliminary; translated from GB/EMBL/DBSJ
A;Molecule type: DNA
A;Residues: 1-202, 'I', 204-237, 'F', 239-460 <LIN>
A;Cross-references: EMBL:AJ22869; NID:G2689196; PIDN:CAA11019.1; PID:G2689197
A;Genetics:
A;Gene: SPC1885.11c; cds1
A;Map position: 3
A;Introns: 86/3; 103/1; 113/3; 142/1; 228/3; 298/3; 368/2
A;Function:
C;Description: EC 2.7.1.-; protein kinase Cds1 [validated, MUID:98119835]; is required for the G2/M transition and activated by S-phase arrest and activated by DNA damage
C;Superfamily: protein kinase Cds1; kinase interaction domain homology; protein kinase h
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;62-133/Domain: kinase interaction domain homology <KH>
F;62-133/Domain: kinase interaction domain homology <KH>
F;165-433/Domain: protein kinase homology <KIN>
F;173-181/Region: protein kinase ATP-binding motif

Query Match 20.5%; Score 77.5; DB 2; Length 460;
Best Local Similarity 30.6%; Pred. No. 0.45;
Matches 22; Conservative 9; Mismatches 34; Indels 7; Gaps 2;

Qy 8 WTGRRGCDLSPFNKLVSGDHCRIV-----VDEKSGQVTLDTSTGTVINKLVVK 61
Db 60 WRFGHKSCVVL-NGPRVSNFPEIYQGHNDSEENVVFLHDHSSNGTFLNFERLAK 118
Qy 62 KQTCLPTQGDVI 73
Db 119 NSRTILNGDEI 130

RESULT 6
T29340
Hypothetical protein F21C10.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29340
R;Du, Z.; Gattung, S.
A;Title: The sequence of C. elegans cosmid F21C10.
A;Reference number: 220610
A;Accession: T29340
A;Status: preliminary; translated from GB/EMBL/DBSJ
A;Molecule type: DNA
A;Residues: 1-2541 <DUZ>
A;Cross-references: EMBL:U55364; PIDN:AAA97973.1; GSPDB:GN00023; CESP:F21C10.7
A;Experimental source: strain Bristol N2; clone F21C10
A;Genetics:
A;Gene: CESP:F21C10.7
A;Map position: 5
A;Introns: 50/3; 91/3; 153/3; 184/3; 215/1; 254/3; 350/1; 392/3; 538/1; 589/3; 625/3; 75

Query Match 20.1%; Score 76; DB 2; Length 2541;
Best Local Similarity 31.0%; Pred. No. 4.3;
Matches 22; Conservative 13; Mismatches 24; Indels 12; Gaps 3;

Qy 5 KREWTIGRRGCDLSPFNKLVSGDHCRIVDE-----KSGQVTLDTSTGTVIN--KLX 58
Db 1265 KEETAVKESERVHLTF-----SGDHQMIIDKTVPLDTGIYTRAKNVHGEVANFCCLR 1318
Qy 59 VVKKQTCPLQT 69
Db 1319 VVPKQPPPT 1329

RESULT 7
SCL protein - mouse (fragment)
B61188
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 13-Mar-1998 #text_change 07-May-1999
C;Accession: B61188
R;Ku, D.H.; Chang, C.; Koniecki, J.; Cannizzaro, L.A.; Boghosian-Sell, L.; Alder, H.; Ba
Cell Growth Differ. 2, 179-186, 1991

A;Title: A new growth-regulated complementary DNA with the sequence of a putative trans-
A;Reference number: A61188; MUID:91329275; PMID:1868030
A;Accession: B61188
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-265 <KUA>

Query Match 19.7%; Score 74.5; DB 2; Length 265;
Best Local Similarity 28.6%; Pred. No. 0.55;
Matches 18; Conservative 18; Mismatches 22; Indels 5; Gaps 3;

Qy 16 CDLSP-PNKK--LVSGDHCRIVDEKSG--QVTLDTSTGTVINKLVKKQTCPLQTG 70
Db 40 CDVALRPOQEPCLISGVHAEFLAEIQGDDWRVSLDEHSSQGTLVNNVRLPGRHRLLESDG 99
Qy 71 DVI 73
Db 100 DIL 102

RESULT 8
S43941
protein kinase DUN1 (EC 2.7.1.-) - Yeast (Saccharomyces cerevisiae)
N;Alternate names: protein D2370; protein YDL101C
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jun-2000
C;Accession: S43941; S67643; S67418; S72106
R;Zhou, Z.; Elledge, S.J.
Cell 75, 1119-1127, 1993
A;Title: DUN1 encodes a protein kinase that controls the DNA damage response in yeast.
A;Reference number: S43941; MUID:94084787; PMID:8261511
A;Accession: S43941
A;Molecule type: DNA
A;Residues: 1-513 <ZHO>
A;Cross-references: EMBL:I25548; NID:G435616; PIDN:AA16324.1; PID:G435617
R;Ballesta, J.P.G.; Remacha, M.; Soler-Mira, A.; Jimenez, A.; Garcia-Cantalejo, J.M.; Boi
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67629
A;Accession: S67643
A;Molecule type: DNA
A;Residues: 1-513 <BAL>
A;Cross-references: EMBL:Z74149; NID:G1431139; PIDN:CAA98668.1; PID:G1431140; GSPDB:GN000
A;Experimental source: strain S288C
R;Bostkovic, J.; Saiz, J.E.; Soler-Mira, A.; Garcia-Cantalejo, J.; Revuelta, J.L.; Jimine
submitted to the EMBL Data Library, February 1996
A;Reference number: S67406
A;Accession: S67418
A;Molecule type: DNA
A;Residues: 1-513 <BOS>
A;Cross-references: EMBL:X95644; NID:G1199535; PIDN:CAA64912.1; PID:G1199548
R;Saiz, J.E.; Buitrago, M.J.; Garcia, R.; Revuelta, J.L.; del Rey, F.
Yeast 12, 1077-1084, 1996
A;Title: The sequence of a 20.3 kb DNA fragment from the left arm of Saccharomyces cerevi
A;Reference number: S72094; MUID:97051597; PMID:8896274
A;Accession: S72106
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-513 <SAI>
A;Cross-references: EMBL:X95644; NID:G1199535; PIDN:CAA64912.1; PID:G1199548
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C;Genetics:
A;Gene: SGD:DUN1; MIPS:YDL101C
A;Cross-references: MIPS:YDL101C; SGD:S0002259
A;Map position: 4L
C;Function:
C;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
C;Superfamily: protein kinase DUN1; kinase interaction domain homology; protein kinase h
C;Keywords: ATP; nucleus; phosphoprotein; phosphotransferase; serine/threonine-specific k
F;58-128/Domain: kinase interaction domain homology <KH>
F;198-480/Domain: protein kinase homology <KIN>
F;206-214/Region: protein kinase ATP-binding motif

Query Match 19.0%; Score 72; DB 1; Length 513;


```

Best Local Similarity 35.7%, Pred. No. 2.2;
Matches 25; Conservative 10; Mismatches 29; Indels 6; Gaps 3;

QY 9 TIGRRGCD--LSFSPSNKLVSGDHCRIVVDEKSGQ---VTLEDSTSTGTVINKLVKVKQ 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 TIGRSRCDVLSBPDISTFAEFPHLLQMDVDFNINVIDKSRNGTFINGNRLVKKD 116
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 64 TCPLQTGDVI 73
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 117 YI-LKNGDRI 125
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
S73738
phosphoglycerate kinase (EC 2.7.2.3) pgk - Mycoplasma pneumoniae (strain ATCC 29342)
N:Alternate names: hypothetical protein A05_orf409
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C:Accession: S73738
R:Hammelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73738
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-409 <HM>
A:Cross-references: EMBL:AE000040; GB:U00089; NID:g1674091; PIDN:AA896060.1; PID:g1674091
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Gene: pgk
A:Genetic code: SGC3
C:Superfamily: phosphoglycerate kinase
C:Keywords: ATP; phosphotransferase

Query Match 18.8%; Score 71; DB 2; Length 409;
Best Local Similarity 29.2%; Pred. No. 2.3;
Matches 19; Conservative 14; Mismatches 20; Indels 12; Gaps 2;

QY 19 SFPNSKLVSGDHCRIV-----VDEKSGQVTLDTSTGTVINKLVKVKQTCPLQ 68
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41 SLDTPIKLLGHNCIVLLSHLRVKSLLDKKKSLOPVASA--LQNLKNTKVFHCPEN 98
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 69 TGDVI 73
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 99 TGDVX 103
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
A83872
hypothetical protein BH1777 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: A83872
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira-
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: A83872
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <STO>
A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA805496.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1777

Query Match 18.7%; Score 70.5; DB 2; Length 230;
Best Local Similarity 31.5%; Pred. No. 1.4;
Matches 23; Conservative 13; Mismatches 30; Indels 7; Gaps 4;

QY 3 LRKREWTIGRRGC---DLSPFNKLVSGDHCRIVVDEKSGQVTLDE-TSTSGTVINKLK 58

```

Query Match 18.0%; Score 68; DB 2; Length 754;
Best Local Similarity 24.7%; Pred. No. 9.8;
Matches 20; Conservative 21; Mismatches 30; Indels 10; Gaps 2;

[illegible]

Search completed: May 7, 2004, 14:50:07
Job time : 5.33553 secs

RESULT 14

G98197

Hypothetical protein AGR_L1057 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C;Species: Agrobacterium tumefaciens

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002

C;Accession: G98197

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qutollo, B.; Goldman, B.

C.; Liu, F.; Woliam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 2323-2328, 2001

A;Title: Genome sequence of the Plant pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: G98197

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-399 <KUR>

A;Cross-references: GB:AE007870; PID:g15158909; GSPDB:GN00170

C;Genetics:

A;Gene: AGR_L1057

A;Map position: linear chromosome

```

Query Match      18.0%; Score 68; DB 2; Length 399;
Best Local Similarity 29.3%; Pred. No. 4.9;
Matches         22; Conservative 16; Mismatches 27; Indels 10; Gaps 4;

QY       7 EWT-----IGRRGCDLSPFSN- KLVSGDHCRIVVDKSGQVTTEDTSTGVINKLK 58
          |||:|||||:|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db        18 QMSFERGREAIQRSDCDQWDDNNRVRYSKLHCTLSRDGE-GFIILL-DQSANGILVDGLR 75

QY       59 VVKVKQTCPLOTGDVI 73

Db       76 LLEGESARLRDGSQI 90

```

RESULT 15
T00393
Nijmegen breakage syndrome protein NBS1 - human